DOCUMENT RESUME

ED 218 311 TM 820 351

AUTHOR Price, Gary G.; And Others
TITLE IGE Evaluation Phase 1 Summary Report: Report from the Program on Evaluation of Practices in Individualized Schooling.
INSTITUTION Wisconsin Center for Education Research, Madison.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
REPORT NO WCER-TR-578
PUB DATE Nov 81
GRANT NIE-G-81-0009
NOTE 124p.; Figures I and II may be marginally legible due to small print.

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS *Academic Achievement; Elementary Education; *Individualized Education Programs; *Instructional Development; Models; *Program Evaluation; School Surveys; Summative Evaluation
IDENTIFIERS *Individually Guided Education; *Instructional Programming Model

ABSTRACT

The relationship of organizational and instructional features in Individually Guided Education (IGE) schools and the achievement of second and fifth grade students was studied in Phase I of the IGE evaluation. A positive relationship between organizational features and teacher job-satisfaction was found in a sample study of 156 schools. Reading and mathematics Instructional and Research (I&R) units in each school were analyzed with student achievement hypothesized to be positively correlated with measures of: (1) the organization of classroom instruction; (2) the use of varied curriculum materials designed to be compatible with the Instructional Programming Model (IPM); (3) the extent of schoolwide implementation; and (4) assorted features distinctive to IGE. None of the hypothesized links with student achievement were supported by the data. The IPM practice of individualizing instructional decisions was found to be facilitated within I&R units and Instructional Improvement Committees. The empirically testable premises of IGE are assessed as an educational system rather than by comparison with other schools. (Author/CM)

ED218311

Technical Report No. 578

IGE Evaluation Phase 1 Summary Report

by Gary G. Price, Thomas A. Romberg, and Terance C. Janicki

TM 820 335

November 1981

Wisconsin Center for Education Research

Technical Report No. 578

IGE EVALUATION PHASE I SUMMARY REPORT

by

Gary G. Price, Thomas A. Romberg and Terance C. Janicki

Report from the Program on
Evaluation of Practices in Individualized Schooling

Gary G. Price and Thomas A. Romberg
Faculty Associates

Wisconsin Center for Education Research
The University of Wisconsin
Madison, Wisconsin

November 1981

The research reported in this paper was funded by the Wisconsin Center for Education Research which is supported in part by a grant from the National Institute of Education (Grant No. NIE-G-81-0009). The opinions expressed in this paper do not necessarily reflect the position, policy, or endorsement of the National Institute of Education.

Wisconsin Center for Education Research

MISSION STATEMENT

The mission of the Wisconsin Research and Development Center is to understand, and to help educators deal with, diversity among students. The Center pursues its mission by conducting and synthesizing research, developing strategies and materials, and disseminating knowledge bearing upon the education of individuals and diverse groups of students in elementary and secondary schools. Specifically, the Center investigates

- diversity as a basic fact of human nature, through studies of learning and development
- diversity as a central challenge for educational techniques, through studies of classroom processes
- diversity as a key issue in relations between individuals and institutions, through studies of school processes
- diversity as a fundamental question in American social thought, through studies of social policy related to education

The Wisconsin Research and Development Center is a noninstructional department of the University of Wisconsin-Madison School of Education. The Center is supported primarily with funds from the National Institute of Education.

Table of Contents

6

List of Tables

List of Figures

Abstract

This is the summary of Phase I of the Evaluation Practices in IGE schools. Phase I was a large sample survey of 156 schools. Two Instruction and Research (I & R) units per school were studied--one that included grade 2 children and one that included grade 5 children. The following were all hypothesized to be positively correlated with measures of student achievement in reading and mathematics: measures of the organization of classroom instruction, a measure of the use of varied curriculum materials designed to be compatible with the Instructional Programming Model (IPM), a measure of the extent of schoolwide implementation of the IPM, and measures of assorted schoolwide organizational features distinctive to IGE. There were also hypothesized links among some of these measures. Although some of the hypothesized links among IGE-related measures were empirically supported, none of the hypothesized links with student achievement were supported by the data. Some of these negative findings can plausibly be attributed to measurement error and specification error inherent to the survey methods used in Phase I; others cannot. Some organizational features distinctive to IGE were found to be positively correlated with teachers' job satisfaction.

The IGE Evaluation Project

This is a summary report on schooling practices of a sample of schools which were identified as implementors of Individually Guided Education (IGE). It is the summary of Phase I, which is one of five related phases of an extensive study designed to evaluate IGE. In the following pages, the purpose and design of the IGE Evaluation Project are described.

Through the combined efforts of the Wisconsin Research and Development Center for Individualized Schooling (R&D Center), the University of Wisconsin IGE Teacher Education Project, the Kettering Foundation (/I/D/E/A/), and IGE coordinators in 25 states, more than 2,000 elementary schools adopted a system of elementary education called Individually Guided Education (IGE). However, prior to the IGE Evaluation Project, no comprehensive picture existed which showed how extensively or how effectively IGE had been implemented in these schools.

Thus, the purpose of the IGE Evaluation Project was to evaluate IGE in order to gain a more comprehensive view of the system's operation and effectiveness. The desired outcome was to identify which features of individualized schooling contribute most to the success of reading and mathematics instruction. It was assumed that the features we would identify could be related to the R&D Center's emerging theory of schooling. In addition, methodological procedures and paradigms used in the study would be of general interest to other scholars studying schooling practices.

For more than a decade, IGE was the primary program of the R&D Center. It is a complex system based on theoretic and pragmatic ideas about schooling,

children's learning, and the professional roles of school staffs. The IGE program contains seven components:

1. Multiunit elementary school organization,
2. Instructional programming for the individual student,
3. Assessment and evaluation for educational decision making,
4. Curriculum and instructional materials and activities for each child's instructional program,
5. Home-school-community relations programs,
6. Facilitative environments for professional growth, and
7. Continuing research and development for system improvement.

Each of these seven components is the result of a long, collaborative study by various scholars and professional educators.

Four types of variables were identified to guide the evaluation of IGE: pupil and staff outcomes, means of instruction, support systems, and pupil and staff background. With these types of variables in mind, a descriptive framework was developed that considers outcomes of IGE as a function of both the means of instruction and the degree of implementation (Romberg, 1976). Figure 1 shows a general a priori framework of how the four types of variables are related.

1. <u>Pupil and staff outcomes, and the extent to which these outcomes have been attained, should be the initial basis of an IGE evaluation.</u> As Klausmeier (1977) stated:

> Students, upon completing IGE elementary schooling, should have achieved more than in other kinds of schools, should have acquired higher-level conceptualizing skills and other abilities which enable them to continue to learn, and also should have developed healthy self-concepts. (p. 7)

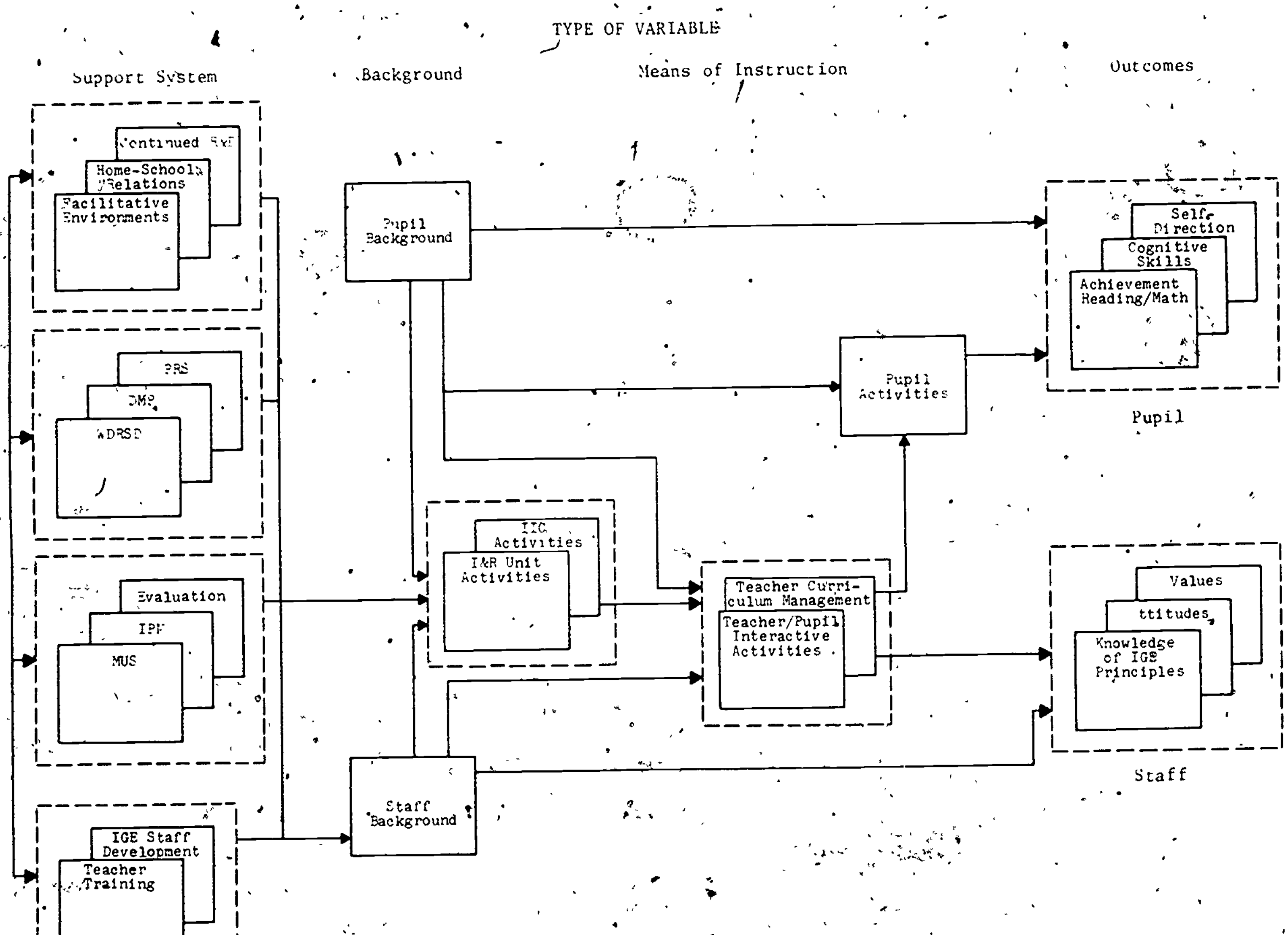


Figure 1. Framework for an IGE evaluation.

Presumably, changes in pupil outcomes are, in part, a result of changed teacher activities. For that reason, any plan for the evaluation of IGE should give attention to staff outcomes--including increases in staff knowledge and changes in attitudes and beliefs.

Both pupil and staff outcomes are illustrated in Figure 1 as being multivariate and multilevel. Pupil outcomes include achievement in reading and mathematics, cognitive skills such as conceptualizing and problem solving, and self-direction. Staff outcomes include attitudes about children and schooling, beliefs about the nature and value of education, and knowledge of IGE principles pertaining to individual differences among children and use of the Instructional Programming Model.

2. The instructional means of formal schooling must be a second basis of an IGE evaluation. It has been fashionable in evaluation circles to concentrate on ends or outcomes and to ignore the means by which they are reached. Reform movements, such as IGE, invariably involve criticisms of and changes in means. Consequently, judging the value of the means is as important as assessing outcomes.

Means of instruction were separated into three sets of activities based upon the operating characteristics of IGE schools: (1) staff activities of the Instructional Improvement Committee (IIC) and the Instruction and Research Units (I & R Units), (2) activities of the staff teacher (in both curriculum management and pupil interactions), and (3) activities of pupils as they are related to reading and mathematics instruction.

3. The degree to which the support systems of IGE have been incorporated and developed in a school must be judged. The seven components of IGE have

evolved as features of schools intended to support new instructional methods and thereby to produce desired pupil and staff outcomes. It can be argued that the effectiveness of an IGE school depends upon which components of IGE have been implemented and how well those components are operating.

The support systems for an IGE learning environment were separated into four categories in Figure 1. The first includes the multiunit organization, instructional programming, and evaluation (IGE Components 1, 2, and 3). The second category, curricular materials compatible with instructional programming and evaluation (IGE Component 4), is represented (but not exhaustively covered) by the three major curricular products developed for IGE. These are the Wisconsin Design for Reading Skill Development (WDRSD) (Otto, 1977), Developing Mathematical Processes (DMP) (Romberg, 1977), and the PreReading Skills Program (PRS) (Venezky & Pittelman, 1977). The third category of support systems includes home-school-community relations, facilitative environments, and continued research and development (Components 5, 6, and 7). The final category of support systems includes the teacher training and IGE staff development programs.

4. The fourth type of variable in Figure 1 is pupil and staff background. Background variables were included because knowledge of prior pupil achievement, level of motivation, and learning style is assumed to be necessary for efficient grouping of students and selection of appropriate activities. Similarly, staff experience with IGE principles, with working in groups, and with the pupils of the particular school is important.

The functional relationships illustrated in Figure 1 convey the following premises: (a) the degree to which IGE support systems have been implemented, together with pupil and staff backgrounds, directly influences the means of

instruction in an IGE school; and (b) the means of instruction, along with pupil and staff backgrounds, account for pupil and staff outcomes.

To clarify how this Phase I report fits into the overall study, all five phases of the project are described below.

Phase I

Phase I consists of a large sample study which provides basic information about IGE schooling. Certain features of IGE schooling have been reputed to be crucial to IGE success. The purpose of Phase I was to examine the extent to which those presumably essential features have been implemented among IGE schools and to assess the effectiveness of that implementation. In this large sample study (including approximately 155 IGE schools), information was gathered using self-report surveys from IGE school staff members and standard paper-and-pencil instruments from students. The data were intended to provide a functional understanding of IGE features, processes, and outcomes by relating a broad scope of variables in an interpretive manner. Using structural equations, Phase I simultaneously examines relationships among the network of variables believed to influence means of instruction, staff outcomes, and pupil outcomes--thereby advancing our understanding of each IGE feature and the network as a whole.

Phase II

Phase II verifies the self-report data gathered in Phase I and extends data collection to include a fuller range of variables. As verification of the Center's Phase I work, Phase II was conducted by Research Triangle Insti-

tute, Roderick A. Ironside, principal investigator, under a subcontract from the R&D Center. The final report is available as Technical Report No. 499 (Ironside & Conaway, 1979). Specifically, the objectives of Phase II were:

1. to determine the validity of the self-report data gathered in Phase I.

2. to use interview and observation data to extend the information about each category of variable.

3. to ascertain the role of developmental agencies in the national diffusion process as perceived in IGE schools.

4. to gather cost data so that some indications of cost-effectiveness could be determined.

The areas of cost and of implementation history, including the role of developmental agencies, are the primary additions of Phase II to Phase I data. The importance of cost analysis has been discussed by Rossmiller and Geske (1977). Adoption and institutionalization of innovative practices are processes which interest not only practitioners and scholars, but also funding agencies (Berman & McLaughlin, 1976).

Phase III

Phase III focuses on the social meanings which emerge as IGE is used on a day-to-day basis. This phase seeks to increase understanding of the impact of educational reform by viewing schools as having a "culture" which governs their members. According to this approach, a school is a complex social arrangement with underlying patterns of conduct that channel the thought and action of persons in the school. These underlying cultural patterns should not be ignored. Failure of educational planners to consider the cultural patterns of schools has produced unanticipated and unintended results from

reform efforts. Research into the impact of educational reform suggests that changes in school programs frequently do not alter existing educational patterns. Instead, the reforms are transformed to fit into the everyday patterns of school life. Early in the development of IGE, the R&D Center explicitly stated that the purpose of IGE schooling was to alter the substantive nature of curriculum and instruction in elementary schools (Klausmeier, Morrow, & Walter, 1968). It was argued that, as schools incorporated IGE, new patterns of conduct, values, and attitudes would emerge. In part, IGE was expected to instill a sense of intellectual vitality, critical thought, and personal responsibility into students' work. From a different perspective, IGE was expected to infuse the staff with a sense of professional participation--an active searching for new information and new skills. But, does the implementation of IGE in fact alter teachers' conceptions of their jobs and students' conceptions of learning? Do new values, attitudes, and patterns of conduct develop in IGE schools, or are old underlying patterns disguised in new forms? Phase III sought answers to these questions and thereby helped to provide a comprehensive perspective for understanding the impact of IGE upon schooling.

Phase IV

Phase IV was designed to evaluate the effectiveness of WDRSD, DMP, and PRS as aids to instructional programming and to examine the relationships of means of instruction and instructional time to pupil outcomes. The success of IGE depends heavily on the availability of curricular materials compatible with instructional programming for the individual student and the availability of appropriate evaluative procedures (Klausmeier, Rossmiller, & Saily, 1977). The three programs developed at the R&D Center--the Wisconsin Design for

Reading Skill Development (WDRSD), Developing Mathematical Processes (DMP), and PreReading Skills Program (PRS)--were designed to meet this need.

Phase V

Phase V synthesizes the results of Phases I through IV of the IGE Evaluation. Each phase of the evaluation study was designed to complement and strengthen the validity of the data gathered by the previous phases. For example, data on means of instruction, gathered by the large-sample study of Phase I, are examined in somewhat greater depth in fewer schools by the Phase II studies. Phase III's analysis develops a view of instruction from a different perspective. Phase IV explores means of instruction within specific curricular areas.

Instead of merely adding together summaries of the different evaluation phases, Phase V proposes to integrate and interpret the data from all the phases into a series of statements about the implications of the project for contemporary educational issues. This phase explores several major themes about schooling by using data from all phases of the evaluation.

I

Introduction

This document provides a discussion of the relationships in Individually Guided Education (IGE) schools between the organizational and instructional features of schools and the reading and math achievement of second and fifth grade students. The document also discusses the relationship between organizational features of IGE and teacher job-satisfaction. The document presents data collected by the large sample survey component of the IGE evaluation, referred to hereafter as Phase I. The theoretical positions that are part of the IGE model--some explicit and some implicit but evident--imply what some of the relationships should be. Thus, we were able to develop a theoretical model of how variations in different organizational and instructional features would be linked, and how, in turn, those variations would be linked to variations in achievement. This report presents that theoretical model and evaluates how adequately it fits data obtained from schools.

The purpose of Phase I, as with the other IGE evaluation phases, is to gain a more comprehensive view of the model's operation and effectiveness. The basic objective is to identify features of IGE schooling that contribute to successful instruction, especially in reading and mathematics. The identification of such features is critical to evaluating IGE as an educational system; it is also integral to understanding schooling in general.

Certain features of IGE schooling have a reputation as keys to the program's success. Phase I examines variations in the extent to which these presumably essential features have been implemented among IGE schools, and assesses how influential such variations in implementation are. As a large sample study

(including approximately 155 IGE schools). Phase I provides basic information about IGE features, processes, and outcomes by relating a broad scope of variables in an interpretive manner. The variables are measures of organizational features, instructional features, student achievement, staff job satisfaction, and other characteristics of students and staff. The specific purposes of Phase I are:

1. to determine the degree to which the seven IGE components have been implemented in IGE schools;

2. to describe and examine the relationship between the implementation of IGE components and means of instruction, particularly in reading and mathematics;

3. to describe and examine the relationship between the implementation of IGE components and staff outcomes; and

4. to describe and examine the relationship--presumably by way of the means of instruction--between the implementation of IGE components and pupil outcomes. This includes such outcomes as reading and mathematics achievement, selected cognitive skills, and aspects of personality development.

The means of instruction and outcome variables of this study are, without question, influenced by multiple, simultaneous causes that resist easy description. Since causal relationships are easier to study in isolation, most research has examined only one or a few of these relationships. Simple comparisons between IGE as an undifferentiated package and other educational alternatives provide us with little information about specific features and processes that occur in IGE schools. Therefore, Phase I simultaneously examines relationships among the network of variables believed to influence means of

instruction, staff outcomes, and pupil outcomes--thereby advancing understanding of each IGE feature and the network of variables as a whole.

The Phase I study was not designed as a direct comparison between IGE schools and other schools. Instead, it is an assessment of certain fundamental and empirically testable premises on which IGE is based. IGE evolved as and was disseminated as an educational system with a distinctive approach to instruction (succinctly described in the Instructional Programming Model), distinctive ways of coordinating the instructional efforts of teachers, distinctive forms for involving teachers in decisions that affect the school as a whole, and a supporting network of other IGE schools, regional coordinators, and disseminators. Some curriculum packages were designed to lend themselves to use of the Instructional Programming Model. IGE as its developers intended it was not an isolated innovation but a complete system. As a complete system, IGE is built upon theoretical positions about the goals of education, the effects of certain forms of instruction, the effects of school organization on instruction, and the effects of linkages that go beyond the walls of the school building.

Some of the premises on which IGE is based are explicit. Others are implicit, but evident. Three general kinds of premises were recognized in the Phase I study. Those are:

1. The instructional practices associated with the Instructional Programming Model make high student achievement more likely (Klausmeier, Karges, & Krupa, 1977, pp. 333-334).

2. Certain systems of record keeping and information collection make it more likely that the instructional practices of the Instructional Programming

Model will occur. Likewise, the use of curriculum materials that lend themselves to record keeping, information collection, and segmentation of curriculum units make those instructional practices more likely to occur.

3. Certain features of school organization make use of the Instructional Programming Model more likely. Those organizational features also make it more likely that the staff will be satisfied with their jobs. The influence of organizational features on job satisfaction is considered at the end of this report.

It should be clear from this discussion that the IGE system of elementary education was based on a set of premises about supportive systems, instructional means, and the impact of those upon both pupil and staff outcomes. In designing the evaluation plan, a descriptive framework relating the various aspects of IGE was proposed (Figure 1). From this framework, a more specific model was developed and variables were identified for aspects of the model, as described in the following sections.

Premises of the First Kind

The first kind of premise includes two instructional practices and their reputed influence on student achievement. The first practice was characterized in the Phase I analysis by a variable named Management of Grouping and Instructional Continuity, the second by a variable named Individualization of Instructional Decisions. These variables and others were scaled from information drawn from several questionnaires treated as one large pool of potentially relevant items.

The IGE model suggests that the organization of classroom instruction should have an effect on student achievement. The Instructional Programming

Model (IPM) is used to group students according to their individual needs, making teacher-student instructional interactions more effective. The variable Management of Grouping and Instructional Continuity (IE) measures several facets of the classroom instructional environment and should, according to the IGE model, be directly related to student learning.

The variable Individualization of Instructional Decisions (IDM) measures the extent to which instructional decisions take into account the individual needs of the student. Taking individual student needs into account is the cornerstone of the IGE model, and it is supposed to be related to student achievement. The individualization of instructional decisions is also integral to the forms of instructional grouping mentioned before.

The postulated causal links between these variables and measures of student achievement are shown in Figure 2. The figure expresses an IGE theory of the causal relationships among the variables. It is meant to express an IGE theory, but it does not always use IGE nomenclature.

Student achievement is the dependent variable in ~~Figure 2~~. Only achievement in reading and mathematics as measured by the California Test of Basic Skills (CTBS) are considered in this report. Other areas of achievement (e.g., science, social studies) could also have been assessed, but because reading and mathematics were the two curricular areas in which materials had been prepared for use in IGE schools, and because most schools in the sample indicated they were implementing the IGE Instructional Programming Model in those curricular areas, the decision to limit achievement assessment to those areas seemed warranted.

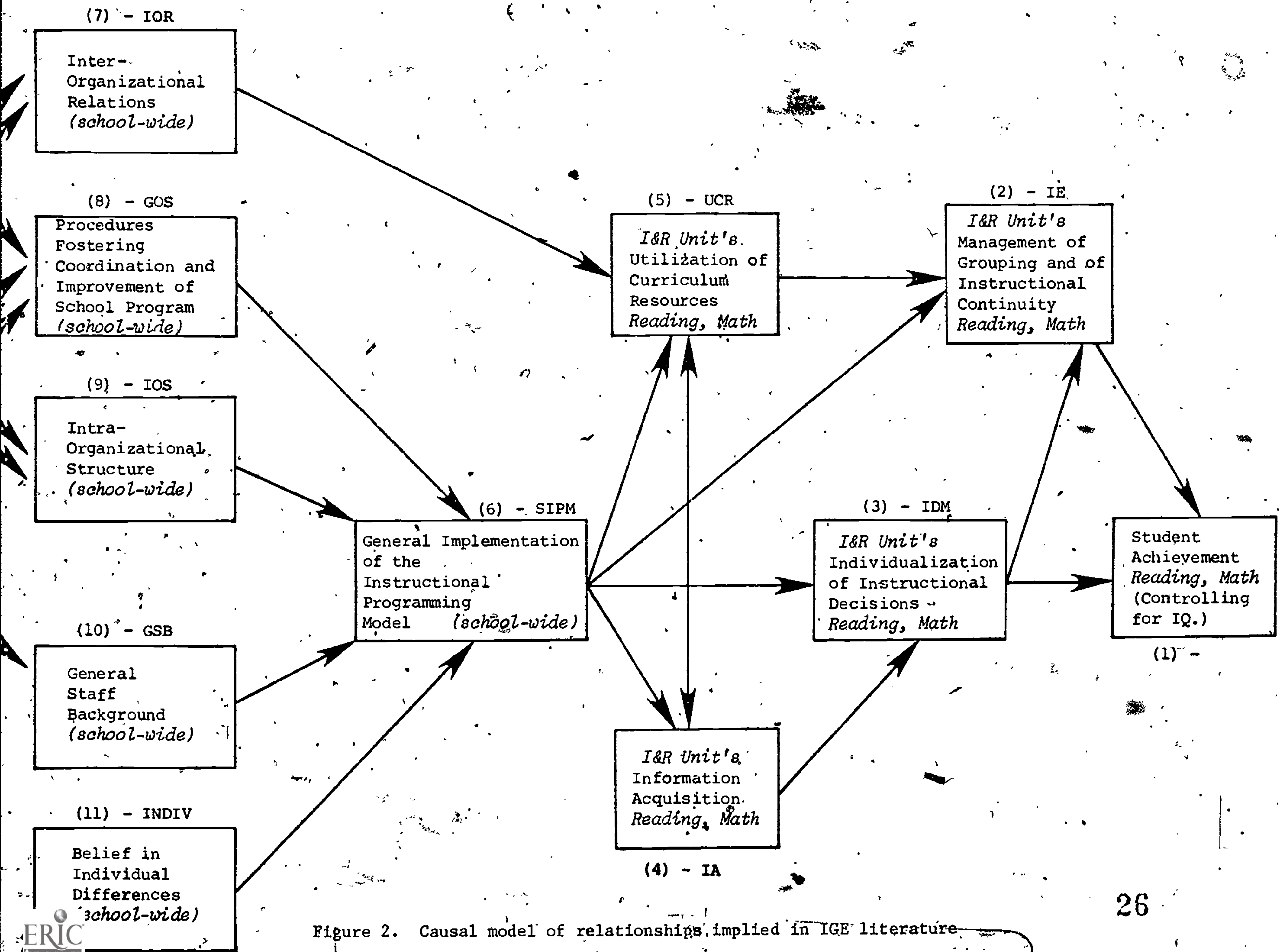


Figure 2. Causal model of relationships implied in IGE literature

Premises of the Second Kind

The second kind of premise involves procedures and features intended to facilitate desirable instructional practices.

Three variables reputed to affect Management of Grouping and Instructional Continuity are shown in Figure 2. One of them has already been mentioned--Individualization of Instructional Decisions. The second is the use of a variety of curriculum resources (UCR), and the third is the schoolwide implementation of the Instructional Programming Model (SIPM). Schoolwide implementation is important because, presumably, the instructional practices of the individual I & R units resonate with the practices elsewhere in the school.

Factors influencing the individualization of instructional decisions. The IGE model suggests that the amount and variety of student information available should directly influence the individualization of instructional decisions. The amount and variety of student information available is reflected by the variable Information Acquisition (IA). The model also suggests that schoolwide implementation of the IPM should promote such individualization. The procedures needed for efficient and useful information acquisition are presumably limited by incomplete schoolwide implementation of the IPM.

Premises of the Third Kind

The third kind of premise involves features of school organization that are distinctive to IGE. These are organizational features intended to make use of the Instructional Programming Model more likely.

Factors influencing general implementation of the IPM. Two organizational features in particular have been presented in the IGE literature as ones which create an organizational environment conducive to the IPM. The first of these

is an assortment of activities collectively labeled and measured as Procedures Fostering Coordination and Improvement of the School Program (GOS). The second, reflected by the variable Intra-Organizational Structure (IOS), is a collection of structural arrangements distinctive to IGE schools, such as organization into I & R units, existence of an Instructional Improvement Committee, and so forth.

The effect of organizational features on schoolwide implementation of the IPM would be difficult to assess if teachers' backgrounds and beliefs were not considered, too. Two appropriate staff measures are included. The first, General Staff Background (GSB), is an aggregate measure of how much IGE-related experience teachers have. The second, Belief in Individual Differences (INDIV), is a measure of how strongly and unanimously the teachers of a school endorse a basic assumption of IGE--the assumption that students differ in ways that instruction ought to take into account.

Factors influencing the use of curriculum resources. One distinctive organizational feature of IGE is the system of linkages between IGE schools. One intended consequence of such Inter-Organizational Relations (IOR) is the exchange of information about IPM-compatible curriculum materials and other curriculum resources. Contact with other schools would presumably help teachers to use a variety of curriculum materials. The variety of curriculum materials used by an I & R unit would depend partly on other practices of the unit, too. For instance, I & R units that gather and organize information about students, through their efforts to individualize instructional decisions, would be more likely to use a variety of curriculum materials.

Summary of the Model

Each arrow shown in the diagram signifies a causal link assumed in the IGE system. Not every causal link assumed in IGE is indicated, however. Suffice it to say that the evaluation staff worked hard to create a model that expresses some of the major causal assertions implicit in IGE. The intent of the evaluation effort was not to ask whether IGE works. Rather, the intent was to ask whether IGE works in the way its developers thought it would. Critics of this approach to evaluation might charge that it accepts the definitions of the program being evaluated. On the other hand, the approach can potentially demonstrate that the factual claims on which a program is based are not borne out empirically. The most obviously serious criticisms are those that meet a model on its own empirical grounds.

The Phase I study evaluates premises of the kind listed above by empirically assessing the relationships they imply. This approach to evaluation would not be feasible if IGE schools had uniformly implemented the organizational features, curriculum features, and instructional practices suggested by experts in IGE. Such uniformity, however, does not exist, as is apparent from the distributions of variables. Phase III of the evaluation, a field study of a few highly reputed IGE schools, provides corroborating evidence that IGE schools are not uniform. In Phase III, marked variability was found even among schools reputed to be exemplary IGE schools.

For the purposes of this study, the fundamental premises listed above have been represented as a network of postulated causal links among the variables of the study. Figure 2 presents these causal links in diagram form. It shows the paths of influence assumed to underlie the relationships between reading achievement and the other variables studied. The figure expresses an IGE

theory of how each variable is causally related to the other variables. The formulation and logic of this model are discussed in the method section.

The diagram is known as a structural model, structural diagram, path diagram, or causal model and follows certain graphic conventions. According to these conventions, a straight, unidirectional arrow signifies that the variable at its base directly influences the variable at its tip. The omission of an arrow constitutes an explicit theoretical statement that no direct causal relationship exists. A variable is an indirect cause of a dependent variable if a path through two or more arrows can be traced from the dependent variable back to the first variable. Associated with each straight arrow is a nonzero value. The sign of the value denotes whether an increase in the causal variable produces an increase (plus) or a decrease (minus) in the dependent variable. A curved, bidirectional arrow is used at the left of the figure between variables which are known to be correlated, but for reasons not covered in the scope of the model.

A system of structural equations corresponds to the model. These equations may be statistically examined for their agreement with the data collected in the Phase I study. Within the limits imposed by measurement error in the group-administered, standardized tests and questionnaires used to collect the data, this approach tests the theoretical model that underlies IGE. If the relationships between variables are not consistent with Figure 2, then probable inaccuracies in the underlying IGE model will be revealed. On the positive side, this approach can indicate important features and processes which deserve more attention from schools implementing IGE. In this context, "importance" means that a feature or process influences outcomes that are socially valued.

II

Methods

Sampling

The population. The population with which this evaluation is concerned is constrained in several ways. The population is necessarily limited to those schools that define themselves as IGE schools. There were between 2,000 and 3,000 such schools in 1974-75 (Klausmeier, 1977). The population studied by this evaluation is further constrained to include only those schools that responded fully to the IGE schools questionnaire of March 1976. There were 946 such schools, 768 of which had both second- and fifth-grade students. The evaluation was limited to students in second and fifth grades, and their teachers. Second grade is the first primary grade in which group-administered paper and pencil tests are sufficiently reliable. Fifth grade is the last elementary school grade common to all elementary schools. This population of 768 IGE schools is nationwide; it includes urban areas, rural areas, low-income areas, and high-income areas.

Sample size. The structural equations analysis is a complex one. Relationships between many variables are explicitly and implicitly assessed. An analysis like this requires many degrees of freedom, hence many schools. In light of the planned analysis, the Phase I IGE evaluation staff decided to seek a sample of 300 schools.

Sampling plan. To select the sample of 300 schools, a plan of stratified random sampling was chosen, in preference to simple random sampling. In heterogeneous populations, where the strata reflect relevant dimensions of heterogeneity, stratified random sampling better ensures that small samples will represent the full variability of the population. The use of stratified random sampling,

however, presupposes that all members of the population (in this case 768 schools) have been classified into mutually exclusive strata. These strata, as implied above, should each have memberships that are relatively homogeneous with respect to the characteristics whose representativeness is being ensured.

Stratification variables. Information about each of the 768 schools in the population was obtained through the March 1976 IGE Schools Questionnaire. That questionnaire concerned the implementation of various IGE features thought to be influential with respect to pupil outcomes, teacher outcomes, and aspects of the instructional process. That information was used to construct seven variables. The seven variables are: (1) rating the staff organization, (2) age of the program, (3) utilization of the Instructional Programming Model, (4) rating for facilitative environment, (5) rating of the organization of children in the school, (6) use of R&D Center curriculum products, and (7) demographic information. These seven variables were then used to classify the population into strata.

The crossing of different dimensions, each of which has been segmented, creates cells containing cases which have received similar values on the variables involved. The use of nonredundant (orthogonal) dimensions provides the most parsimonious means of classifying schools into groups that are similar with respect to multiple, correlated variables. Since the stratification-relevant variables contained redundant information, the variables were orthogonalized using principal components analysis (Harman, 1967). It was decided that 64 strata (cells) was the largest useful number, resulting in approximately 11 schools per stratum. This permitted the population to be segmented into quartiles along the first principal component and median splits (50th percentile) on four other principal components.

The reason for creating a sampling frame was to ensure that the sample would reflect fully the range of variability on the characteristics measured by the seven stratification variables. Selection from such a sampling frame can be systematic, or it can be random (stratified random sampling). To ensure that schools were chosen from the sampling frame without bias, the selection was random and 302 schools were selected. The sample on which the evaluation was ultimately based, however, was not this initially drawn sample. Far fewer schools than the IGE evaluation team had anticipated agreed to participate in the evaluation.

There was consequently a danger that self-selection had created a sample that differed significantly from the population in terms of the IGE characteristics with which the evaluation is concerned. In the absence of a carefully developed sampling frame, we would have been unable to determine whether the participating schools differed from the population (or from the initially selected sample). The sampling frame, however, provided a means of assessing whether and how the participant schools differ from the population. Remarkably, the self-selected sample appears very similar to the population in terms of the stratification variables. (For a more detailed account of the sampling procedure see Price, 1977.)

Staff Instruments

Questionnaires for the staff component of the evaluation were based primarily on several existing instruments which were modified in content or format to meet the requirements of IGE terminology and certain technical constraints, such as machine-readable response forms. A detailed discussion of the content and source of each instrument as well as its relationship to the variables of interest in Phase I appears in Stewart (1977).

The instruments were printed in three questionnaire booklets for distribution according to respondent group: Instructional Improvement Committee and principal, grade 2 and grade 5 test units, and all members of the professional staff. One instrument, the Verification Copy of the IGE Schools Questionnaire, was provided separately for the principal alone. Machine-readable response forms were developed to accompany each questionnaire booklet. A booklet, response form(s), and pencil were packed in an individual envelope preprinted with identification numbers and directions for distribution, completion, and return by the three respondent types.

The time required for an individual staff member to complete the questionnaires varied according to the number of respondent groups to which he or she belonged. Since the unit leaders of the grade 2 and grade 5 test units were usually members of all three respondent groups, they needed to invest the greatest amount of time--245 minutes. The estimated working time for each booklet was:

Instructional Improvement Committee Questionnaire	25 minutes
General Staff Questionnaire	100 minutes
Unit Questionnaire	120 minutes

General staff questionnaires. Every professional staff member assigned to the school at least one-third time during both the 1976-77 and 1977-78 school years was asked to complete the general questionnaires. By setting these criteria, it was expected that all support personnel, such as guidance counselors and remedial teachers, would then participate in the evaluation, but that staff not yet familiar with IGE or not an integral part of building operations would not do so. Aides were not included because of the problems associated with involving instructional aides but not clerical aides.

The five questionnaires combined in the single staff booklet and estimated working times were:

Staff Background Information	20 minutes
IGE Implementation Survey	30 minutes
Job Satisfaction Survey	20 minutes
Assumptions About Conditions of Effective Schooling	15 minutes
Assumptions About Learning and Knowledge	15 minutes
	100 minutes total

On the Staff Background Information Form (Mendenhall, 1977) each staff member provided data on his or her assignment and position in the school, professional and IGE-related training and experience, professional activities, and personal information.

Responses to the IGE Implementation Survey (Klausmeier, 1976) indicated each person's judgment of local implementation of seven IGE components on a five-point scale from no implementation to ideal implementation. There were 77 statements or concepts in the survey, each related to one of the following components: Multiunit Organizational Administrative Arrangements (MUS), Instructional Programming for the Individual Student (IPM), Curricular Programs (CURR), Evaluation for Decision-Making (EVAL), Home-School-Community Relations (HSC), Facilitative Environments (ENV), and Continuing Research and Development (R&D).

The Job Satisfaction Survey (Mendenhall, 1977) measured eight aspects of job satisfaction on a five-point scale ranging from <u>not satisfied</u> to <u>very satisfied</u>. The eight subscales represented in the 50 items were: co-workers,

career function, school identification, financial aspects, working conditions, pupil relations, community relations, and administration.

Staff members were asked to indicate the relative need for each condition listed in a set of 28 statements called Assumptions About Conditions of Effective Schooling. The statements were derived from IGE literature which defined desirable conditions for teaching and learning (Lipham & Fruth, 1976). Responses were on a five-point scale from strongly disagree to strongly agree.

On the last instrument in the General Staff Questionnaire booklet, Assumptions About Learning and Knowledge (Barth, 1971), school personnel reported their feelings about a set of 28 statements related to open education. Response choices again ranged from strongly disagree to strongly agree on a five-point scale.

Unit staff questionnaires. Staff evaluation materials for both Instruction and Research (I & R) units, grades 2 and 5, were identical. However, to insure that the two I & R units responded separately, the covers of the unit questionnaire booklets were printed in different colors, and both the covers and envelopes indicated grade 2 or grade 5.

All members of each I & R unit, except aides, were requested to respond to the questionnaires if they met the following criteria: (1) assignment to the unit at least one-third time in both the 1976-77 and 1977-78 school years, and (2) direct involvement in planning instruction for unit students in reading and/or mathematics and/or language arts. It was assumed that these criteria would allow and encourage part-time unit members such as reading specialists or learning disability teachers to respond and yet avoid participation by personnel unfamiliar with unit operations. Aides were not included.

The three instruments in the unit questionnaire booklet and the estimated working times were:

Role of the Staff Teacher	10 minutes
Instructional Practices in Reading, Mathematics, and Language Arts	60 minutes
Instruction and Research Unit Structure and Function	15 minutes
	85 minutes total

Each unit staff member responded independently to the first two instruments, while the unit staff as a group was asked to complete the third instrument.

Role of the Staff Teacher (Ironside, 1972) pertained to the number and kind of instructional and advisor contact teachers have with students and also to teacher specialties. Instructional Practices in Reading, Mathematics, and Language Arts (adopted by T. J. Fox from DeVault, 1973) assessed four aspects of individualized instruction--rate, media, grouping, and learner assessment--separately for each of the three content areas. The Instruction and Research Unit Structure and Function Instrument (Ironside, 1972) included unit membership, meeting schedules and reports, utilization of meeting time, and the role of instructional and clerical aides. Also incorporated in this questionnaire were questions on topics of special concern in the evaluation: record-keeping and grouping/regrouping practices, time allocations by curricular area, and number of instructional objectives.

Instructional Improvement Committee questionnaires. If an IIC was operating in the school, members were asked to complete by consensus the Instructional (Program) Improvement Committee Structure and Function Questionnaire (Ironside, 1972). Topics in this questionnaire were membership, meeting schedule and

reports of meetings, time allocation to various tasks, and development of schoolwide instructional objectives in the various curricular areas. If no IIC existed or was operative in the school, the principal was asked to describe in brief narrative form the governing body or leadership group or person and to complete the section of schoolwide objectives. The IIC booklet also included a question regarding the demographic category of the school attendance area. Categories were based on the community types developed by the National Assessment of Educational Progress project.

The verification copy of the IGE Schools Questionnaire was distributed with the IIC materials. Principals were asked to verify or update the information previously provided; of particular importance were changes in the schools' organization and extent of application of the Instructional Programming Model, for example, addition of another subject area in which the IPM is applied. The estimated time required for completing both IIC/principal questionnaires was 25 minutes.

Student Instruments

Across the 162 Phase I schools from which student data were received, approximately 5,500 students at grade 2 and 5,800 students at grade 5 were tested. Most schools tested between 24 and 50 students at each grade. These students were preassigned by the R&D Center to one of four test groups (A, B, C, or D) at each grade. Also preassigned were the tests to be administered to each group. The Short Form Test of Academic Aptitude (SFTAA) was the only test administered to all groups. Since the total time requirement was approximately 90 minutes for grade 2 students and approximately 120 minutes for grade 5 students, the tests were scheduled to be administered in three separate sessions.

Listed in Table 1 are the tests administered at both grades. Test copies of the SFTAA, California Test of Basic Skills (CTBS), and Self-Observation Scales (SOS) were purchased from the commercial publisher; machine-scorable SFTAA and CTBS test books were used at grade 2, while reusable test books for these two measures were used with a combination SFTAA-CTBS answer sheet for grade 5. Since each CTBS Reading and Mathematics test was assigned to a specific group, students used only a portion of a test book or answer sheet.

The SOS is a self-report, group-administered instrument with empirically determined scales which measure the way children perceive themselves and their relationships to peers, teacher, and school. The Primary Level measures four dimensions of children's self concept: self-acceptance, social maturity, school affiliation, and self-security. Seven dimensions are measured with the Intermediate Level: self-acceptance, self-security, social maturity, social confidence, school affiliation, teacher affiliation, and peer affiliation.

The five Concept Attainment Abilities (CAA) tests are from a battery of tests that was developed as a part of a previously completed Center project (Harris & Harris, 1973). Only students in grade 5 participated in that study. The CAA tests were administered to assess student cognitive skills in three categories--numerical ability, memory, and word fluency. The tests correspond to these categories as follows:

Cognitive Ability	Test
Numerical	Number Series
	Number Relations
Memory	Picture Class Memory
	Remembering Classes: Members

Table 1

Phase I Student Tests

Test	Grade(s)
Short Form Test of Academic Aptitude (SFTAA)[a]	2,5
Comprehensive Tests of Basic Skills (CTBS), Form S[b]	
Reading Vocabulary	2,5
Reading Comprehension: Sentences	2
Reading Comprehension: Passages	2
Reading Comprehension	5
Mathematics Computation	2,5
Mathematics Concepts & Applications	2,5
Spelling[c]	5
Self Observation Scales (SOS),[d] Form C	2,5
Locus of Control (Cromwell, 1964)[e]	2,5
Concept Attainment Abilities (CAA)	
Number Series	5
Number Relations	5
Picture Class Memory	5
Remembering Classes: Members	5
Omelet	5

[a]Grade 2, Level 1; Grade 5, Level 3.

[b]Grade 2, Level C; Grade 5, Level 2.

[c]Administered as a CAA word fluency substitute.

[d]Grade 2, Primary Level; Grade 5, Intermediate Level.

[e]Not included in the analysis because of low reliability.

Word fluency	Omelet (an anagrams test)
	Spelling (CTBS)

Since comparable tests are not available for grade 2, the CAA tests were given only at grade 5 for the Phase I evaluation.

Measures of School and I & R Unit Characteristics

The structural model was formed before the variables used in the model were created. However, the model was formed with knowledge of the information from which the variables would be created. The initial model was developed by the evaluation staff to satisfy the following conditions: (a) to be a fair rendering of the IGE model; (b) to be specifiable in terms of the item pool; and (c) to have only "main effect" variables (i.e., no attempt was made to address interactive relations among variables).

First, all items from all nonstudent instruments were grouped together. Second, the evaluative staff independently divided the items into sets of items each representative of a single variable. Having followed this procedure independently, members of the group then met and reached consensus on groups of items that defined a particular variable. After this procedure for selecting item sets, each group of items was given a verbal description that reflected the information contained in its constituents. Next the selected items in each group were combined to form a composite variable.

To appraise how well a model fits data, one must begin with trustworthy data. A weakness of the Phase I data stems from the fact that already existing questionnaires were used. Had there been resources and time, it clearly would have been preferable to build questionnaire items specifically to measure the constructs of an a priori model. As it was, available

items were used to scale the variables. For several variables, however, there were many pertinent items available, so, for those variables, little damage was done by relying upon available questionnaires.

A second limitation of the data stems from the remoteness of the data collection procedure and the amount of time it took school staff members to respond to the questionnaires. This questionnaire form of data collection invites hurried responding. Moreover, respondents may have wanted to respond as they thought good IGE citizens should, since they knew that the responses would be sent back to the Wisconsin Research & Development Center. There is no doubt that the signal one receives comes partly veiled with noise when one uses mailed questionnaires.

The evaluation staff was aware of these limitations from the outset and took steps to minimize their effect. An innovative use of questionnaire responses was developed which deserves mention. Usually, questionnaire items, like test items, are combined in a linear manner to form a scale. That was not always done in our study. Rather, Boolean logical expressions were often used to combine the responses on several items into new, composite items to be arranged as a scale. For measurement purposes, these composites were not themselves present in any questionnaire. They were the product of logical operations performed on questionnaire items.

There were two reasons for taking this approach to scaling. One was to "goof-proof" our variables. The detection of contradictions and other convergent uses of questionnaire responses were used to minimize the extent to which our scaled values could be thrown off by erroneous responses. Some forms of deliberate distortion were anticipated, and scaling decisions were made to minimize their effect. A second reason for the approach had nothing to do

with accuracy of responses. In some cases we decided that, even if we assumed the responses were perfectly accurate, a justifiable ordering would not be obtained by arithmetic combination of item responses.

As the report from a field validation study (Ironside & Conaway, 1979) showed, there was ample reason for attempting to "goof-proof" the variables. In that study, Phase II of the evaluation, which involved 30 of the Phase I schools, a field check was run on the validity of questionnaire responses. Correlations between Phase II field ratings and Phase I questionnaire-derived values are available for some variables:

Interorganizational Relations	.66
Procedures Fostering Coordination and Improvement of the School Program	.53
Intraorganizational Structure	.60
General Implementation of the IPM	.67

Those variables, concerned as they are were surface organizational features of IGE schools, were probably measured better than the instructional practices.

III

Description of Variables

Eight of the variables used in this study measure aspects of the school as a whole, such as organizational features, schoolwide practices, staff background, demography, and teacher job satisfaction. Other variables do not pertain to the school as a whole; they are measures of a specific I & R unit's practices in reading and in math. Each variable specific to an I & R unit is actually treated as four separate variables. It is measured in two I & R units per school--one that includes children of grade 2 age and one that includes children of grade 5 age. In each I & R unit, there is a reading version and a math version of the variable. The schoolwide variables are described first. Variables specific to the I & R unit are described thereafter.

Schoolwide Variables

Seven variables describe general characteristics of schools and their staffs. Three of the schoolwide variables, Interorganizational Relations (IOR), Intraorganizational Structure (IOS), and Procedures Fostering Coordination and Improvement of the School Program (GOS), are measures of organizational features. These organizational features are supposed to foster staff job satisfaction and effective instructional practices. Two other schoolwide variables, General Staff Background (GSB) and Belief in Individual Differences (INDIV), are measures of teachers' preparation, experience, and beliefs. These characteristics should also be related to job satisfaction and the use of effective instructional practices. Our structural model makes no formal attempt to enumerate underlying factors that influence differences among schools on the

five variables mentioned thus far. For the purposes of the Phase I study, differences among schools on these variables are taken as given. Accordingly, Figure 2 shows no straight arrows pointing toward these five variables. In the technical vocabulary of structural equations methods, such variables are called <u>exogenous variables</u>.

A sixth general schoolwide variable is Schoolwide Implementation of the Instructional Programming Model (SIPM). Differences among schools on this variable are thought to be affected by four of the exogenous variables (namely, GOS, IOS, GSB, and INDIV). From the point of view of IGE implementers, IGE's distinctive organizational structures, schoolwide procedures, and interorganizational relations are meant to facilitate implementation of the Instructional Programming Model. SIPM is not an exogenous variable, because our model enumerates expected causes of its variations. Variables whose determinants are specified by a model, like SIPM, are referred to as <u>endogenous variables</u>.

A seventh schoolwide variable is Teacher Job Satisfaction (JOBSAT). Job satisfaction--itself an outcome of interest--is kept out of the analysis that deals with student achievement. Results concerning JOBSAT are given later in the report.

The eighth variable, School's Demographic Setting (DB--standing for demographic background), describes the community and student population served by a school. It is not supposed to be related to the other schoolwide variables, nor should it be related to general aspects of instructional practices in the I & R unit. DB has been included because of its well known correlation with student achievement.

In the following paragraphs, each of these schoolwide variables is given a verbal definition. A description of each variable's distribution and internal consistency is also given. Readers wanting to know more about a specific variable should refer to the detailed technical report which corresponds to that variable; Table 2 lists technical reports and the variables with which they correspond. Table 3 presents summary statistics for these variables.

The relation of internal consistency to reliability is such that a composite whose constituents have high internal consistency may or may not possess high test-retest reliability. Conversely, a composite that has high test-retest reliability may have low internal consistency. Any composite that measures a homogeneous construct should be expected to have high internal consistency. However, a composite that measures a construct combining several empirically unrelated facets would not necessarily have high internal consistency, as noted by Cronbach and Meehl (1955). Several of the composites used in this study fit the latter description, and, accordingly, should not be expected to have high internal consistency.

Interorganizational Relations (IOR). IOR is a measure of the school's relationships and activities with outside persons and organizations, especially relationships and activities believed to facilitate implementing and maintaining IGE in the school. IOR includes extraorganizational arrangements and activities of the school and its staff members that keep the school informed of community needs and new educational ideas.

Intraorganizational Structure (IOS). IOS is a measure of certain aspects of the school's internal organization that are relevant to implementing IGE. Organizational structures within the school (Instructional Improvement Committee, Instruction & Research units, etc.) are assessed for characteristics

Table 2

Technical Reports Describing School Variables

Used in Phase I Study

Schoolwide Variables	Technical Report No.
Interorganizational Relations (IOR)	476[a]
Procedures Fostering Coordination and Improvement of the School Program (GOS)	477[a]
General Staff Background (GSB)	478[a]
Intraorganizational Structure (IOS)	479[a]
Belief in Individual Differences (INDIV)	480[a]
School's Demographic Setting (DB)	482[a]
Schoolwide Implementation of the Instructional Programming Model (SIPM)	483[a]
Teacher Job Satisfaction (JOBSAT)	484[a], 512[b]

[a]Price, Janicki, Howard, Stewart, Buchanan, & Romberg (1978a - 1978h).

[b]Price, Janicki, VanDeventer, & Romberg (1979).

Table 3

Summary Statistics for Schoolwide Variables

Variable	Mean	S.D.	Median	Actual Minimum	Logical Minimum	Actual Maximum	Logical Maximum	α
IOR	20.486	6.809	20.228	6.286	0	39.812	42	.64
IOS	20.522	3.802	21.223	8.092	1.50	2[illegible]15	33.3	.63
GOS	58.178	9.908	57.934	27.083	0	83.682	104	.67
GSB	3.844	.618	3.896	1.884	0	5.021	7.0	.26
INDIV	3.163	.179	3.152	2.556	0	3.667	4.0	.87
SIPM	62.441	12.263	64.061	16.578	0	93.446	100	.87
DB	4.096	1.502	4.225	1	1	7	7	—[a]
JOBSAT	40.571	5.079	41.001	24.600	0	50.000	50	.94

[a] DB is a one item variable; therefore no internal consistency can be calculated.

such as membership composition, frequency of meetings, permanence of leadership, amount of release time made available for meetings, whether parents and others participate in the activities of such groups, whether agenda of meetings are kept, and how agenda are distributed. The existence and responsibilities of certain supplementary staff positions (IMC directors, student teachers, aides, and interns) are also assessed as part of the internal organization of the school.

Procedures Fostering Coordination and Improvement of the School Program (GOS). GOS is a measure of procedures in the school that are supposed to foster continuing improvement of the overall school program. Included are research and development, staff development, use of volunteers and aides, noninstructional (advisory) contact between teachers and students, and other aspects of home-school-community relations. We regret that the abbreviated name of this measure, GOS, has little mnemonic value. "GOS" stems from an earlier name by which we knew this variable--retained for the benefit of readers who will go on to read Technical Report 477 (Price et al., 1978b) where "GOS" appears in many computing statements.

General Staff Background (GSB). GSB is a measure of teaching experience in IGE, teaching experience in general, and leadership activities in the profession.

Belief in Individual Differences (INDIV). INDIV is a measure of teachers' belief that students have individual needs which should be considered in planning and implementing an instructional program.

General, Schoolwide Implementation of the Instructional Programming Model (SIPM). SIPM is a measure of implementation of general school practices that have been encouraged by the Wisconsin R&D Center as supportive of the Instruc-

tional Programming Model (IPM). SIPM is based on self-reported practices of: (a) setting schoolwide instructional objectives by the Instructional Improvement Committee (IIC); (b) adapting schoolwide objectives in Instruction & Research (I & R) units; (c) using IIC guidance in the development of recordkeeping procedures; and (d) providing for carrying out the IPM in the I & R units of the school.

<u>School's Demographic Setting (DB)</u>. DB is an ordinal demographic classification of the community served by a school. As such, it is a proxy for experiences, skills, and attitudes that students acquire outside the school, but which affect outcomes of interest to educators. This measure classifies the community of a school into one of the seven demographic categories employed in the National Assessment of Educational Progress; the categories are arranged in order of the rank they had on student achievement in the National Assessment. DB was given consideration as a covariate for analyses involving student achievement. It turned out, however, that DB was negligibly correlated with measures of organizational features and school practices. Its multiple correlation with the six schoolwide variables just described was .197, which is not statistically significant [$F(6,149) = 1.01$, $p > .40$]. Its effect as a covariate was nil, so it was not employed as a covariate in the analyses reported here.

<u>Teacher Job Satisfaction (JOBSAT)</u>. JOBSAT is a schoolwide measure of teachers' satisfaction with a variety of aspects concerning their school and professions. It was based on the 50-item questionnaire that is contained in Technical Report 512 (Price, Janicki, VanDeventer, & Romberg, 1979).

Measures of I & R Units' Practices in Reading and in Math

There are four variables that measure aspects of instruction in a specific I & R unit, rather than in the school as a whole. These variables are Utilization of Curriculum Resources (UCR), Information Acquisition (IA), Individualization of Instructional Decisions (IDM), and Management of Grouping and Instructional Continuity (IE). The practices of an I & R unit can differ from one curriculum area to another. Consequently, these variables have been measured separately in two curriculum areas, reading and math, which are the areas for which measures of student achievement have been obtained in this study. These four variables are intended to measure processes that are closer to what pupils actually experience than the schoolwide variables--processes which are supposed to mediate the influence of schoolwide variables.

Two I & R units from each school are included in the study--one that includes children of grade 2 age and one that includes children of grade 5 age. Each measure of I & R unit practices may thus be thought of as having four versions: practices of the grade 2 I & R unit in reading, the grade 2 I & R unit in math, the grade 5 I & R unit in reading, and the grade 5 I & R unit in math.

In the following paragraphs, each measure of I & R unit practices is given a verbal definition. Readers wanting to know more about a specific variable should refer to the detailed technical report which corresponds to that variable. Table 4 lists technical reports and the variables to which they correspond. Summary statistics for the four versions of each variable are given in Table 5.

Table 4

Technical Reports Describing Unit

Variables Used in Phase I Study

I & R Unit Specific Variable	Technical Report No.[a]
Utilization of Curriculum Resources (UCR)	485
Information Acquisition (IA)	486
Individualization of Instructional Decisions (IDM)	487
Management of Grouping and of Instructional Continuity (IE)	488

[a]Price, Janicki, Howard, Stewart, Buchanan, & Romberg (1978i - 1978l).

Table 5

Summary Statistics for Unit Variables

Version	N	Mean	SD	Median	Actual minimum	Logical minimum	Actual maximum	Logical maximum	α
Utilization of Curriculum Resources (UCR)									
Grade 2, Math	151	39.528	9.611	39.500	11.944	0	60	60	.29
Grade 2, Reading	156	44.845	9.587	46.453	22.000	0	60	60	.27
Grade 5, Math	151	40.148	9.530	40.000	12.000	0	60	60	.20
Grade 5, Reading	156	44.770	9.890	44.667	19.910	0	60	60	.35
Information Acquisition (IA)									
Grade 2, Math	151	7.176	1.276	8.000	4.0	1	8.0	8	.00
Grade 2, Reading	156	7.005	.997	7.001	4.0	1	8.0	8	.04
Grade 5, Math	151	7.386	1.155	8.000	3.0	1	8.0	8	.07
Grade 5, Reading	156	6.928	1.015	7.000	4.5	1	8.0	8	.00

Table 5 (continued)

Version	N	Mean	SD	Median	Actual minimum	Logical minimum	Actual maximum	Logical maximum	α
Individualization of Instructional Decisions (IDM)									
Grade 2, Math	151	8.332	2.605	8.000	1.000	0	14.00	14	.02
Grade 2, Reading	156	8.158	2.379	8.223	2.333	0	13.00	14	.08
Grade 5, Math	151	8.092	2.553	8.000	1.000	0	14.00	14	.10
Grade 5, Reading	156	7.511	2.181	7.125	1.500	0	13.00	14	.04
Management of Grouping and of Instructional Continuity (IE)									
Grade 2, Math	151	29.581	7.258	30.000	14.000	0	45.750	48	.19
Grade 2, Reading	156	28.638	6.120	28.338	13.668	0	45.000	48	.58
Grade 5, Math	151	28.536	7.805	28.250	12.000	0	45.000	48	.22
Grade 5, Reading	156	26.339	5.907	26.000	11.000	0	42.875	48	.63

Utilization of Curriculum Resources (UCR). UCR is a measure of the extent to which the I & R unit uses a variety of IPM-compatible curricular materials and media in a particular subject area (reading, math).

Information Acquisition (IA). IA is a measure of: (a) the completeness of information gathered for making decisions about the instruction of individual students; (b) the variety of means for gathering such information; and (c) the use of record keeping procedures by the I & R unit for organizing and retrieving the information. The measure is specific to subject area and to each I & R unit.

Individualization of Instructional Decisions (IDM). IDM is a measure of the extent to which decisions affecting instruction are adapted to differences among students. This is made evident by the basis used for moving children to new material, the different rates of individual progress through curricular units, and the percentage of I & R unit meeting time spent making instructional decisions.

Management of Grouping and of Instructional Continuity (IE). IE is a measure of several aspects of environmental organizations. This variable includes arrangement of the instructional environment in terms of scheduling, group size and change patterns, and opportunities for children to become aware of interrelationships and continuity in instruction.

Measures of Student Achievement

It should be noted that the final analysis of pupil data involved only the achievement measures for reading and mathematics. Data on student cognitive abilities and personality development had been gathered, but were not used in the analysis.

Grade 2. The Comprehensive Tests of Basic Skills (CTBS), Form S, Level C, Expanded Edition, provided measures of reading and mathematics achievement in Grade 2. The following subtests of the CTBS were administered: reading vocabulary, reading comprehension of sentences, reading comprehension of passages, mathematics computation, and mathematics concepts and applications. Because of time constraints, most students took only some of the CTBS subtests. It was unnecessary to require each student to take every test, since the grade 2 instruction and research (I & R) unit, rather than the individual student, was the unit of analysis. CTBS subtest scores of individual pupils in each I & R unit were averaged to form an aggregate measure of the I & R unit. Prior to that aggregation, however, adjustments were made to individuals' subtest scores. Those adjustments are described in the next section of this report and, in more detail, in Technical Report 408 (Price, VanDeventer, Janicki, & Romberg, 1979.)

Grade 5. The same general approach taken in grade 2 I & R units to form overall reading achievement and mathematics achievement scores was also taken in grade 5 I & R units. A different level of achievement test was used, however, and that test included some additional subtests. In grade 5, the Comprehensive Tests of Basic Skills (CTBS), Form S, Level 2, provided measures of reading and mathematics achievement. The following subtests were administered: reading vocabulary, reading comprehension, mathematics computation, and mathematics concepts and applications. More detail about grade 5 achievement scores and their adjustment was given in Technical Report 509 (Price, VanDeventer, Janicki, & Romberg, 1979h).

IV

Adjustment and Aggregation of Student Achievement Scores

Literature promoting IGE suggests that student achievement will be raised by implementing the features of IGE (e.g., Klausmeier, 1977; Klausmeier, Karges, & Krupa, 1977). Therefore, student achievement measures are an important part of the Phase I evaluation. One would expect schools high on implementing IGE features to have high achieving students. Examining the relations between IGE features and student achievement, however, can be misleading if differences due to other influences are not accounted for. Students from a school in a wealthy suburb, for example, would commonly score higher on an achievement test than students from a school in a poor urban neighborhood (Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld, & York, 1966). Some differences in academic achievement between rich and poor children appear to persist even when their school programs are identical (Firkowska, Ostrowska, Soskolowska, Stein, Susser, & Wald, 1978). Achievement differences associated with community type cannot be ignored. Presumably, a major reason why students in rich and poor neighborhood schools would differ in achievement is that they are unequal at the outset in their command of the knowledge and skills required for successful school performance.

Schools are unevenly matched in terms of the student populations they serve. That fact is commonly recognized when schools are compared directly. Teachers and administrators from different schools are usually aware of the disparate backgrounds from which each other's students come. However, the need to control for differences in student background is less obvious when relationships between school characteristics and student achievement are being studied. The influence of a school characteristic or practice on achievement

can be obscured if differences in student background are not controlled. An educational practice that has no effect on achievement can be correlated with achievement if most schools using the practice serve a population of educationally advantaged students. Conversely, an effective practice can appear ineffective if most schools using it serve a population of educationally disadvantaged students.

To control for differences in student achievement that result from differences in student populations, we obtained two measures: the demographic and economic character of each school's community (demographic background), and student's academic ability.

Demographic background pertains to the school neighborhood as a whole. Consequently, it is not sensitive to educationally important differences that exist between the backgrounds of children who attend the same school. It simply represents the kind of neighborhood being served: suburban, inner city, small town, rural, etc. As such, demographic background was seen as a reasonable proxy for the school setting, the educational ambience of the community, and the prevailing economic circumstances of students' families. As mentioned earlier, the school demographic background measure proved to be negligibly related to other school characteristics measured. It was correlated with student achievement scores (from .29 with Mathematics achievement in grade 2 to .39 with Reading achievement in grade 5), but its correlation was less once adjustment was made for individual students' scores on the Short Form Test of Academic Aptitude.

The Short Form Test of Academic Aptitude (SFTAA), a measure of scholastically useful skills and knowledge, was given to each child. This measure provided information about individual students' short-term prospects for success in school subjects. A high score on the SFTAA suggests that a student

has acquired--through home, prior schooling, and elsewhere--skills and knowledge that make it easier to succeed in school. Since there are innumerable ways to get a low score on tests such as the SFTAA, the educational prospects of low scoring students are less clear. Nevertheless, a low score suggests that a student may lack skills and knowledge useful to school learning. The SFTAA, therefore, gives probabilistic information about scholastically useful skills and knowledge of students. Like any probabilistic information, it can be wrong in individual cases.

The academic ability scores of individual students can be used to predict how those students will fare on subject area achievement tests. That prediction, in fact, was our primary means of making adjustments for achievement test differences that could not be reasonably attributed to school program differences. However, prediction on the basis of information specific to the individual student reveals only part of the effect academic ability can have on achievement test performance. By aggregating the SFTAA scores of students in a school, a measure of the school's educational milieu is provided. This can augment the prediction provided by individual scores. In other words, one might predict different achievement from two students who obtained the same SFTAA score, if one attended a school in which most students scored low on the SFTAA and the other attended a school in which most scored high.

Student raw scores on achievement measures were carefully adjusted to provide a better comparison among schools. The student achievement variables used in subsequent analyses of the Phase I study were not simply raw scores, but residual scores derived from the differences between raw and predicted scores on achievement tests. Each student's scores on academic ability tests were used to predict his or her scores on the achievement tests. The student's actual scores on those measures were then compared with the predictions. The

deviations of the actual scores from the predicted scores were the variables of interest. These difference scores, known technically as regression residuals, were then aggregated.

Guarding Against Faults in SFTAA Scores

The adjusted achievement scores, central to subsequent analyses, are susceptible to any inaccuracies that might exist in SFTAA scores. An unrepresentative score on the SFTAA would cause an inappropriate adjustment to be made on a student's achievement subtest scores. Therefore, special pains were taken to assure that this basis for adjustment had values which seemed reasonable.

One cause of unrepresentative SFTAA scores is mechanical errors in taking the test. Second grade pupils, who have little experience in taking paper-and-pencil tests, are likely to make such errors.

Faulty test administration is another possible source of difficulty. If a teacher's explanations or directions are unclear, student scores can be adversely affected. Some questions on the SFTAA are read to the students by the teacher, so some teacher-caused problems are inevitable. Errors in timing the tests could distort the scores considerably. Finally, a student might not achieve a representative score because of a bad day, home troubles, illness, test anxiety, discomfort about the testing room, or distractions in a group setting.

Winsorization. A procedure known as "winsorization" was used to help avoid problems that can beset a one-time measure of academic ability. Generally, winsorization involves replacing extreme scores with scores that are still on the outskirts of a distribution, but which are less extreme (see Anscombe, 1960; Dixon & Tukey, 1968). In effect, winsorization pulls improb-

ably high or low scores back to a believable range. Unrepresentative low scores on the SFTAA occur much more frequently than unrepresentative high scores, so only the lower tail of the distribution was winsorized for this study.

Thus far, we have described winsorization as a technique used only to adjust extreme values in a univariate distribution--that is, values of a particular variable that are so extreme as to be implausible. In addition, winsorization can be applied to outliers from a bivariate distribution. Implausible pair-wise combinations of values are singled out for winsorization in a bivariate distribution.

Winsorization with respect to bivariate distribution was the approach used in the present study. The two variables, whose joint distribution was adjusted for outliers, are raw scores from the two halves of the SFTAA: language ability and nonlanguage ability. These subtests are correlated, but it is common for individuals to have a higher normative standing on one subtest than on the other. It is presently fashionable in educational practice to interpret a large discrepancy between such scores as symptomatic of problems that can interfere with school learning. However, a test that is susceptible to misrepresenting a child's underlying competence is a poor basis for making inferences about such problems. Consequently, we make no psychoeducational interpretation of discrepancies between children's language ability and nonlanguage ability scores.

Some children will have authentically large differences between scores on two tests, yet for a particular population of children there are bounds within which discrepancies must stay to be credible. Discrepancies beyond those bounds can be caused by measurement procedures that have gone awry.

Alternatively, they can accurately describe a child who is so atypical that the evaluator should question whether it is taxonomically correct to include the child in the same population as the other children. Children who are very atypical should be removed from an analysis to avoid a distorted picture of the population. If there is insufficient evidence to justify such removal, then winsorization is appropriate. Winsorization does not eliminate large discrepancies, but it reduces them to a plausible level.

As mentioned previously, unrepresentatively low scores were more likely to occur than unrepresentatively high scores. Therefore, when the discrepancy between language ability and nonlanguage ability was implausibly large, the lower of the two scores was winsorized. The higher score was not.

Based on the bivariate distribution, 95% confidence bands for individual scores were calculated for language ability and nonlanguage ability, respectively. To make these calculations, each variable had to be regressed on the other. Scores falling too far below one of the regression lines were adjusted back to the lower bound of the confidence interval for that regression line.

In practice, students with low scores on both language and nonlanguage ability, or with two high scores, did not have their scores adjusted. Only students with relatively good scores on one subtest and very poor scores on the other received adjusted scores. By adjusting only as far as the confidence band, the distortion of erroneous outliers was minimized. Out of 5,352 students, only the extremely discrepant scores of 242 students were affected by the winsorization procedure.

<u>Adjusted Achievement Scores as Regression Residuals</u>

The SFTAA language ability and nonlanguage ability measures were then used as covariates to adjust the CTBS achievement subtest scores of each

student. Scores of all students who took an achievement subtest (N approximately 1,384) were used in the regression analysis. Each CTBS subtest (in raw score form) was regressed on SFTAA language ability and nonlanguage ability (also in raw score form) and on sex.[1]

Each student's predicted score on each subtest was estimated from the regression equation. The predicted score was then subtracted from the student's actual raw score on the subtest, thereby creating a residual score for each subtest. This residual score reflects the student's achievement after being statistically adjusted to control for the effects of academic ability and sex.

Aggregation of Adjusted Achievement Scores

The residual scores of individual students on CTBS subtests provided the basis for an aggregated measure representing each school. The subtests are, in the case of grade 2, reading vocabulary, reading comprehension of sentences, reading comprehension of passages, mathematics computation, and mathematics concepts and applications. The subtests for grade 5 are, in sequence, reading vocabulary, reading comprehension, mathematics computation, mathematics concepts, and mathematics applications. The scores of each group of students in each school were averaged to form a school mean for each subtest. In other words,

---

[1]Performance on the SFTAA can be scaled in several ways. We used raw score scaling in preference to IQ scaling. When an ability test such as the SFTAA is being used to predict school achievement, its IQ scale form can yield unintended results unless age is also included as a predictor.

the adjusted scores (regression residuals) of students in an I & R unit were averaged to provide a representative measure of (ability-adjusted) student achievement in that I & R unit.

After aggregation, I & R unit means on residual scores from achievement subtests were combined to form an overall reading achievement score and an overall math achievement score. Recall that overall scores could not have been formed for individual students because, with a few exceptions, a given individual took only some of the subtests. The overall reading and math achievement scores--formed by simple addition of I & R unit averages of ability-adjusted subtest scores--are the primary achievement measures used in later analyses.

Tables 6 and 7 provide descriptive statistics about the univariate distributions of all ability and achievement measures, using the individual pupil as the unit of analysis. Tables 8 and 9 present descriptive statistics, using the I & R unit as the unit of analysis.

Table 6

Grade 2 Student Ability and Achievement

Descriptive Statistics for Individuals

Variable	Valid N	Mean	Standard deviation	Minimum	Maximum
<u>Short Form Test of Academic Aptitude (SFTAA)</u>					
Language ability (raw score)	5,346	29.688	6.517		
Language ability (adjusted by winsorization)	5,349	29.724	6.436	8.368	45.000
Nonlanguage ability (raw score)	5,312	26.732	6.678		
Nonlanguage ability (adjusted by winsorization)	5,349	26.747	6.583	8.674	40.000
<u>Comprehensive Test of Basic Skills (CTBS)</u>					
Reading vocabulary	1,391	22.694	7.667	4.000	33.000
Reading vocabulary (adjusted)	1,380	- .133	6.210	-22.347	15.158
Reading comprehension of sentences	1,391	14.352	6.939	1.000	23.000
Reading comprehension of sentences (adjusted)	1,380	- .122	6.130	18.135	12.758
Reading comprehension of passages	1,367	11.191	5.539	.000	18.000
Reading comprehension of passages (adjusted)	1,352	- .034	4.616	-13.946	11.643
Mathematics computation	1,376	13.688	5.643	1.000	28.000
Mathematics computation (adjusted)	1,362	- .044	4.924	-12.156	17.047
Mathematics concepts and applications	1,402	15.381	5.072	2.000	25.000
Mathematics concepts and applications (adjusted)	1,390	- .025	3.857	-12.432	13.065

Table 7

Grade 5 Student Ability and Achievement

Descriptive Statistics for Individuals

Variable	Valid *N*	Mean	Standard deviation	Minimum	Maximum
Short Form Test of Academic Aptitude (SFTAA)					
Language ability (raw score)	5,578	21.209	8.520	1.000	45.000
Language ability (adjusted by winsorization)	5,604	21.201	8.501	1.000	45.000
Nonlanguage ability (raw score)	5,586	24.264	6.953	3.000	40.000
Nonlanguage ability (adjusted by winsorization)	5,604	24.272	6.911	4.000	40.000
Comprehensive Test of Basic Skills (CTBS)					
Reading vocabulary	1,477	24.979	8.129	3.000	40.000
Reading vocabulary (adjusted)	1,468	-.118	4.564	-22.974	15.715
Reading comprehension	1,450	27.196	9.523	2.000	45.000
Reading comprehension (adjusted)	1,436	-.045	5.622	-27.797	15.596
Mathematics computation	1,416	28.855	9.022	5.000	47.000
Mathematics computation (adjusted)	1,401	.057	6.832	-21.112	21.073
Mathematics concepts	1,423	15.106	4.461	2.000	25.000
Mathematics concepts (adjusted)	1,398	.026	2.940	-11.802	9.250
Mathematics applications	1,422	14.507	5.818	.000	25.000
Mathematics applications (adjusted)	1,397	-.009	3.921	-17.053	11.807

Table 8

Grade 2 Student Ability and Achievement

Descriptive Statistics for School Aggregates

Variable	Valid N	Mean	Standard deviation	Minimum	Maximum
Short Form Test of Academic Aptitude (SFTAA)					
Language ability (adjusted by winsorization)	156	29.794	2.929	19.708	36.944
Nonlanguage ability (adjusted by winsorization)	156	26.814	2.865	16.253	32.000
Comprehensive Test of Basic Skills (CTBS)					
Reading vocabulary	155	22.683	4.071	8.167	30.500
Reading vocabulary (adjusted)	155	- .168	3.123	-9.521	6.521
Reading comprehension of sentences	155	14.327	3.132	6.000	21.308
Reading comprehension of sentences (adjusted)	155	- .152	2.669	-8.081	5.547
Reading comprehension of passages	155	11.107	2.814	3.429	16.333
Reading comprehension of Passages	155	- .121	2.065	-5.244	5.645
Mathematics computation	155	13.619	3.036	5.625	23.500
Mathematics computation (adjusted)	155	- .121	2.545	-7.355	10.286
Mathematics concepts and applications	154	15.375	2.715	8.571	21.100
Mathematics concepts and applications (adjusted)	154	- .041	1.870	-6.323	3.647
Reading total score	154	48.114	8.985	21.000	65.538
Reading total score (adjusted)	154	- .454	6.687	-18.909	12.402
Mathematics total score	153	29.007	4.869	16.054	43.856
Mathematics total score (adjusted)	153	- .260	3.470	-19.857	7.627

Table 9

Grade 5 Student Ability and Achievement

Descriptive Statistics for School Aggregates

Variable	Valid *N*	Mean	Standard deviation	Minimum	Maximum
Short Form Test of Academic Aptitude (SFTAA)					
Language ability (adjusted by winsorization)	156	21.135	3.317	9.375	29.286
Nonlanguage ability (adjusted by winsorization)	156	24.213	2.879	15.072	30.833
Comprehensive Test of Basic Skills (CTBS)					
Reading vocabulary	156	24.902	4.134	11.818	32.000
Reading vocabulary (adjusted)	156	- .137	1.948	-9.172	4.671
Reading comprehension	155	27.044	4.571	13.000	37.400
Reading comprehension (adjusted)	154	- .117	2.465	-10.685	6.348
Mathematics computation	153	28.834	4.347	15.833	38.125
Mathematics computation (adjusted)	153	.009	3.305	-8.993	11.547
Mathematics concepts	154	15.057	2.274	8.400	20.500
Mathematics concepts (adjusted)	153	- .011	1.332	-3.963	3.072
Mathematics applications	154	14.457	2.914	6.000	20.833
Mathematics applications (adjusted)	153	- .044	1.706	-6.075	4.229
Reading total score	155	51.935	7.655	24.818	67.100
Reading total score (adjusted)	154	- .260	3.470	-19.857	7.627
Mathematics total score	152	58.400	7.808	34.400	76.042
Mathematics total score (adjusted)	152	- .097	4.597	-11.823	14.019

V

Results and Discussion

Reading and Math Achievement

This section examines features of IGE schools believed to affect achievement in reading and math. The positive relationships between features of IGE schools and measures of students' achievement (as implied in the model) were not found. In grade 2 I & R units, no organizational variable and no measure of instructional practices was correlated beyond a trivial level with either reading achievement or math achievement, although two correlations were statistically significant. Reading achievement had a correlation of .197 ($p < .01$) with General Staff Background (GSB), and math achievement had a correlation of .255 ($p < .01$) with General Implementation of the Instructional Programming Model.

The results in grade 5 I & R units were just as negative. Every variable was correlated only trivially with the reading achievement and math achievement variables. Not a single correlation differed from zero with statistical significance. Accordingly, the structural equation analysis lends no empirical support to hypothesized paths of influence.

This negative finding may, in part, indicate faults in our measurement of organizational features and instructional practices. However, not all of the blame can be placed on the attenuating effect of noise in the measures, because ratings of organizational features made by field observers in 30 schools showed reasonably high correlation with the questionnaire-based scales (IOR, .66; GOS, .53; IOS, .60; SIPM, .67). There is less assurance that our measures of instructional practices are trustworthy.

The standardized, group-administered measure of reading achievement can also be criticized as a less-than-perfect indicator of what children know

about reading. The test may have some sections that do not reflect the curriculum of the schools in our study. On the whole, however, the instrument seems to provide a reasonable assessment. Other information we have indicates that the test scores are reasonably accurate. Different demographic classifications of schools, for instance, show differences on the reading achievement and math achievement measures that closely parallel the findings in the National Assessment of Educational Progress.

Organizational Features and Instructional Practices

The following discussion examines links between organizational features and the instructional practices they were meant to facilitate. Analysis began with an estimation of the a priori model. The initial estimation procedure was a straightforward application of multiple regression. The correlation matrix for schoolwide measures is given in Table 10. The correlation matrices of additional variables used in the regression analyses are given in Tables 11 and 12; summary statistics for these variables were given in Tables 3 and 5. Each endogenous variable was regressed on its theorized causes. Endogenous variables, those hypothesized to be under the influence of other variables in the model, have at least one straight arrow pointing toward them in a causal diagram. Exogenous variables are those not under the influence of other variables in the model.

Adjacent to each path in Figures 3, 4, 5, and 6 are path coefficients (standardized multiple regression coefficients) as estimated for the full a priori models of grade 2 reading, grade 2 math, grade 5 reading, and grade 5 math, respectively. These coefficients were subjected to a statistical test. The a priori model was revised on the basis of those tests. Causal links hypothesized in the a priori models whose path coefficients did not differ statistically from zero were dropped from the models. In effect, such a pro-

Table 10

Correlation Matrix for Schoolwide Measures

(N = 156 schools)

	SIPM	IOR	GOS	IOS	GSB	INDIV
SIPM	1.00000					
IOR	.60854	1.00000				
GOS	.72897	.72345	1.00000			
IOS	.49570	.46323	.56988	1.00000		
GSB	.26864	.23079	.24814	.27759	1.00000	
INDIV	.45289	.43155	.47223	.20199	.16266	1.00000

Table 11

Reading Correlation Matrices for Reading Analysis

	IOR	GOS	IOS	GSB	INDIV	SIPM	UCRREAD	IAREAD	IDMREAD	IEREAD	ACH.-READ
					Grade 2						
UCRREAD	.15205	.17846	.12725	.15381	.11666	.21802	1.00000				
IAREAD	.15507	.25631	.11606	.06925	.29210	.31110	.13904	1.00000			
IDMREAD	.18010	.26506	.21058	.05681	.09314	.25806	.24168	.21317	1.00000		
IEREAD	.10159	.13499	.05800	.02489	.16992	.21395	.11377	.15357	.35616	1.00000	
ACH.-READ	-.06695	-.10528	-.07779	.19749	-.11062	-.03360	-.08624	-.04865	.00786	-.01926	1.00000
					Grade 5						
UCRREAD	.17408	.16749	.03755	.15325	.16385	.20324	1.00000				
IAREAD	.15821	.14556	.22917	.07371	.25673	.18552	.10194	1.00000			
IDMREAD	.18543	.13222	.26817	.24024	.25756	.26017	.15875	.33666	1.00000		
IEREAD	.14491	.01421	-.00291	.02329	.17535	.13441	.16371	.21933	.28312	1.00000	
ACH.-READ	.12003	.05247	.01872	.09591	-.14392	-.01230	.06137	-.05028	-.03345	-.07301	1.00000

Note: Correlations presented are by pairwise deletion. For grade 2, 149 ≤ N ≤ 154; for grade 5, 147 ≤ N ≤ 155.

Table 12

Correlation Matrices for Mathematics Analyses

	IOR	GOS	IOS	GSB	INDIV	SIPM	UCRMATH	IAMATH	IDMMATH	IEMATH	ACH.-MATH
Grade 2											
UCRMATH	.05461	.18196	.04341	.08216	.16589	.12369	1.00000				
IAMATH	.18512	.28016	.17778	.10369	.33295	.30899	.12524	1.00000			
IDMMATH	.09937	.23457	.22372	.07899	.18472	.22697	.06024	.16786	1.00000		
IEMATH	.07139	.14405	.14559	.07569	.17832	.18894	.01940	.27100	.32554	1.00000	.00362
ACH.-MATH	.08778	.08633	.06342	.07185	.13485	.25525	-.05999	.03532	.10130	.00362	1.00000
Grade 5											
UCRMATH	.10950	.20235	.02094	.06060	.15819	.15175	1.00000				
IAMATH	.16362	.14873	.11683	.01812	.22176	.08625	.15141	1.00000			
IDMMATH	.10836	.09788	.18035	.20125	.16347	.20342	.11955	.24383	1.00000		
IEMATH	.10236	.02448	.03116	-.05289	.01766	.05719	.20083	.24691	.19905	1.00000	
ACH.-MATH	.13695	-.00995	-.06048	-.01086	-.07241	-.03432	.03183	.02593	-.06153	.01654	1.00000

Note: Correlations presented are by pairwise deletion. For grade 2, $149 \leq N \leq 153$; for grade 5, $147 \leq N \leq 152$.

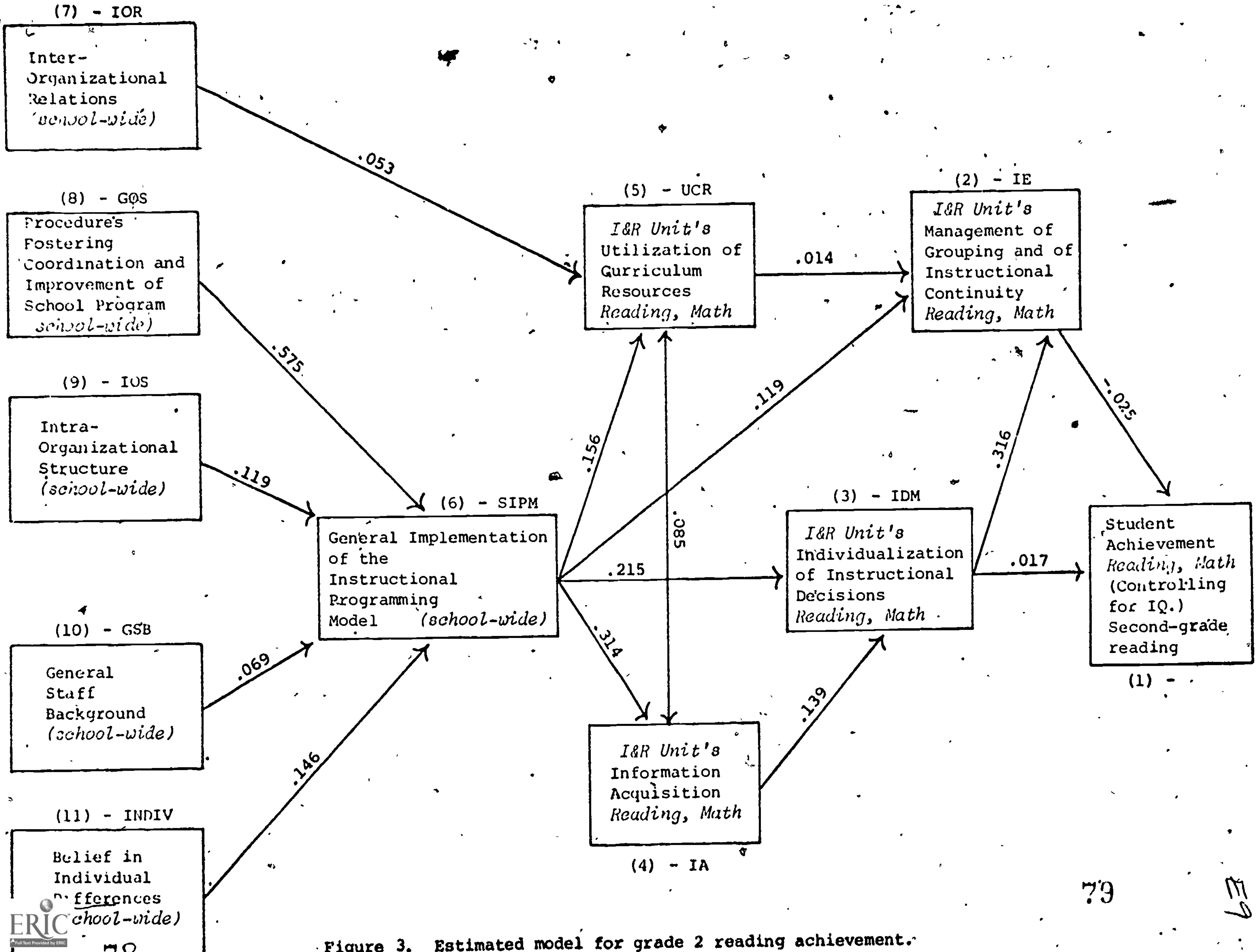


Figure 3. Estimated model for grade 2 reading achievement.

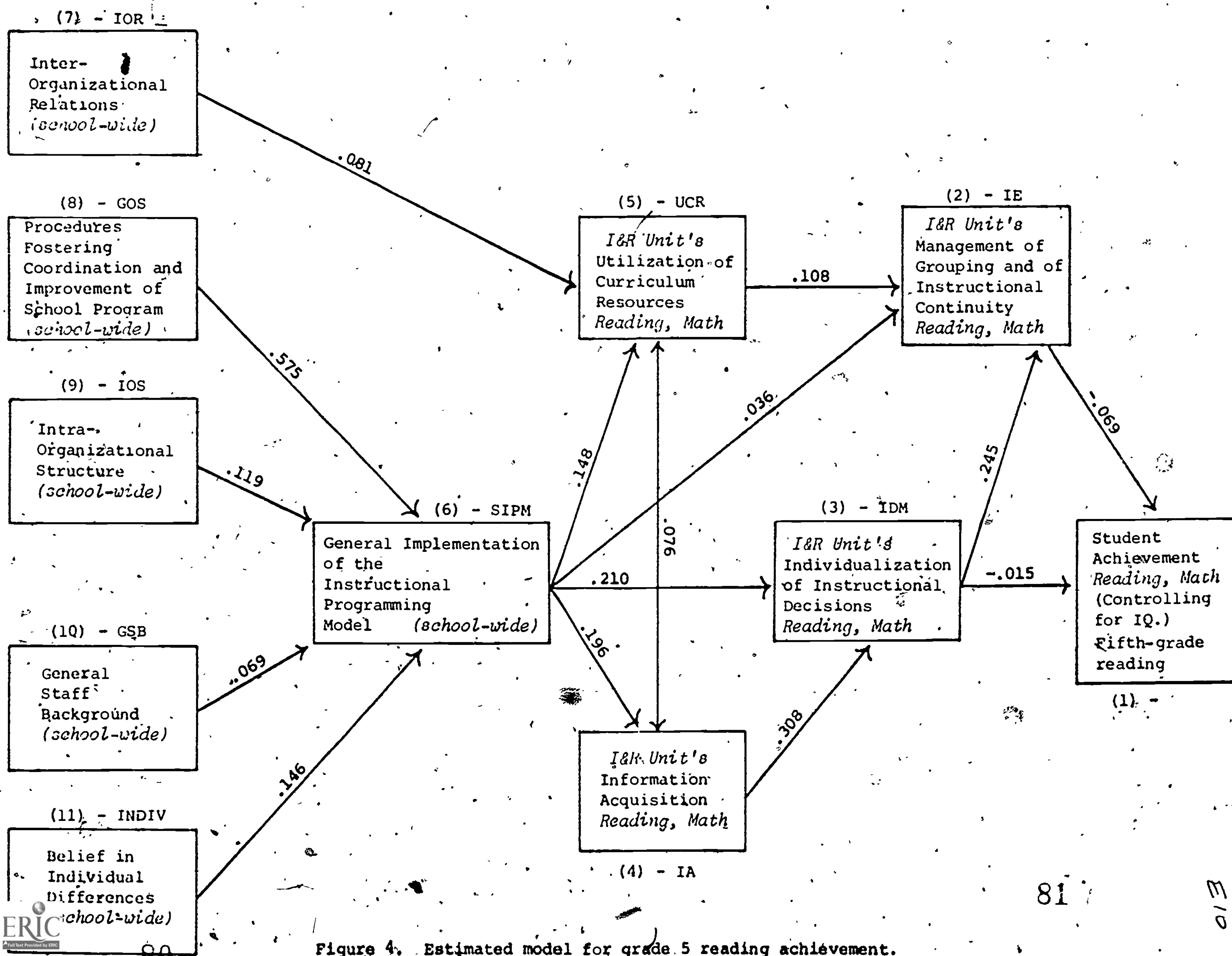


Figure 4. Estimated model for grade 5 reading achievement.

cedure tests whether a variable used as a predictor accounts for any unique variance in the dependent variable.

Paths theoretically precluded from an a priori model were hypothesized to make no increment to $\underline{R}^2$ beyond that obtained by using the paths included in the model. The omitted paths of greatest interest lead directly from remote causes to effects--bypassing theorized mediators. When the addition of a direct path from a remote cause--a path theoretically precluded heretofore--was statistically significant, a model was deemed insufficient to explain the observed relationship between the effect and that remote cause. Figures 7, 8, 9, and 10 show the empirically revised (fitted) models. The information in these figures can also be represented algebraically as structural equations. Structural equations for the hypothesized models and the fitted models are given in Technical Reports 510 and 511 (Price, Janicki, VanDeventer, & Romberg, 1980a, 1980b).

However, when predictors are interrelated--as they are here--interpretation of the tests for particular path coefficients can easily be misleading (see Goldberger, 1964; Gordon, 1968). As Tables 11 and 12 make evident, the organizational features in this study were interrelated, as were the instructional variables; therefore, the substantive interpretation of individual path coefficients becomes ambiguous. For this reason we are being deliberately wary of discussing individual regression coefficients.

To get a clearer understanding of the relationships between the organizational and instructional variables, another type of analysis was used. We used orthogonalized predictors, which are unrelated to each other and are linear combinations of the initial set of predictors. Interpretation of coefficients of orthogonalized predictors is possible only to the extent that the orthogonalized predictors can be defined in a substantively meaningful way.

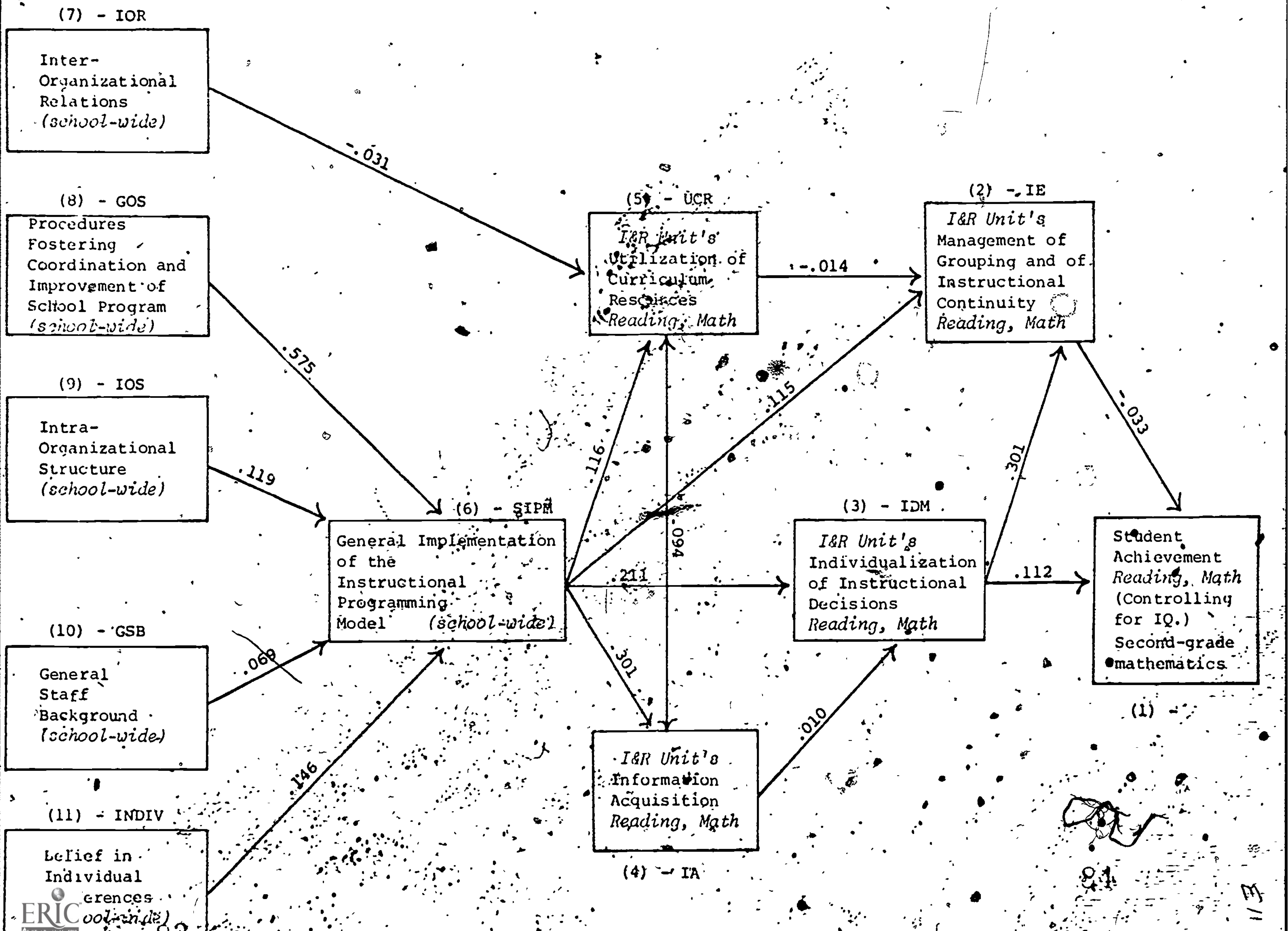


Figure 5. Estimated model for grade 2 mathematics achievement.

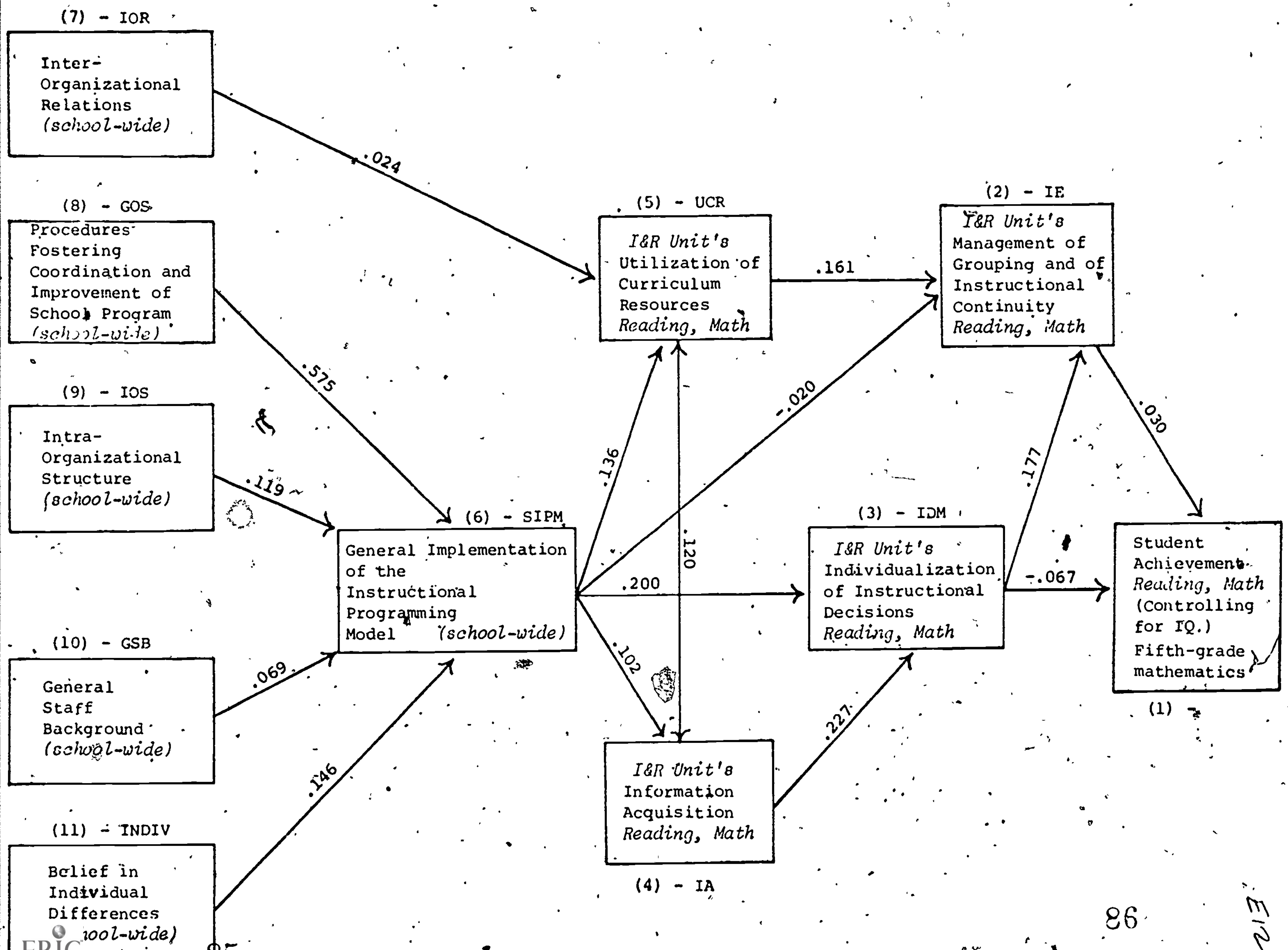


Figure 6. Estimated model for grade 5 mathematics achievement.

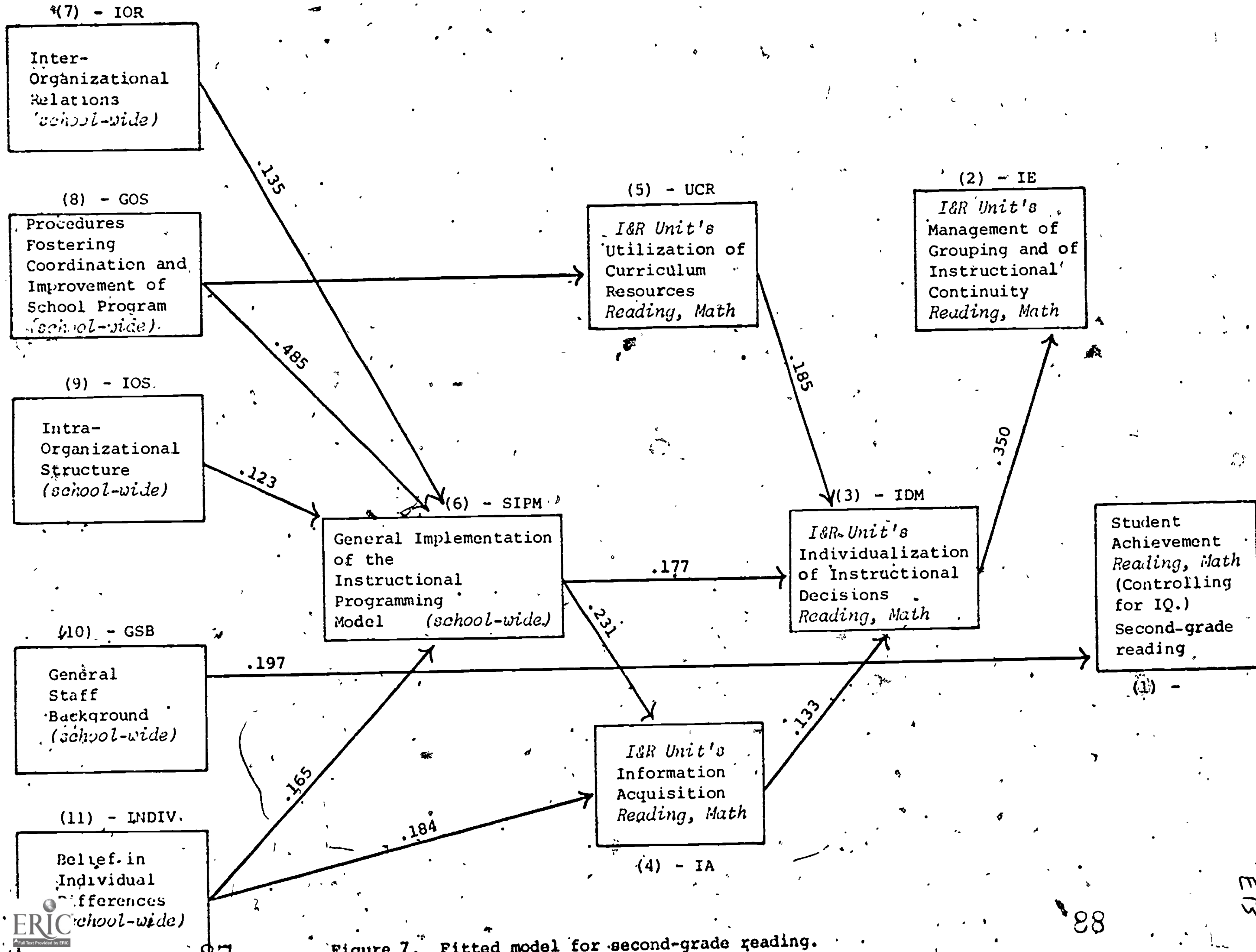
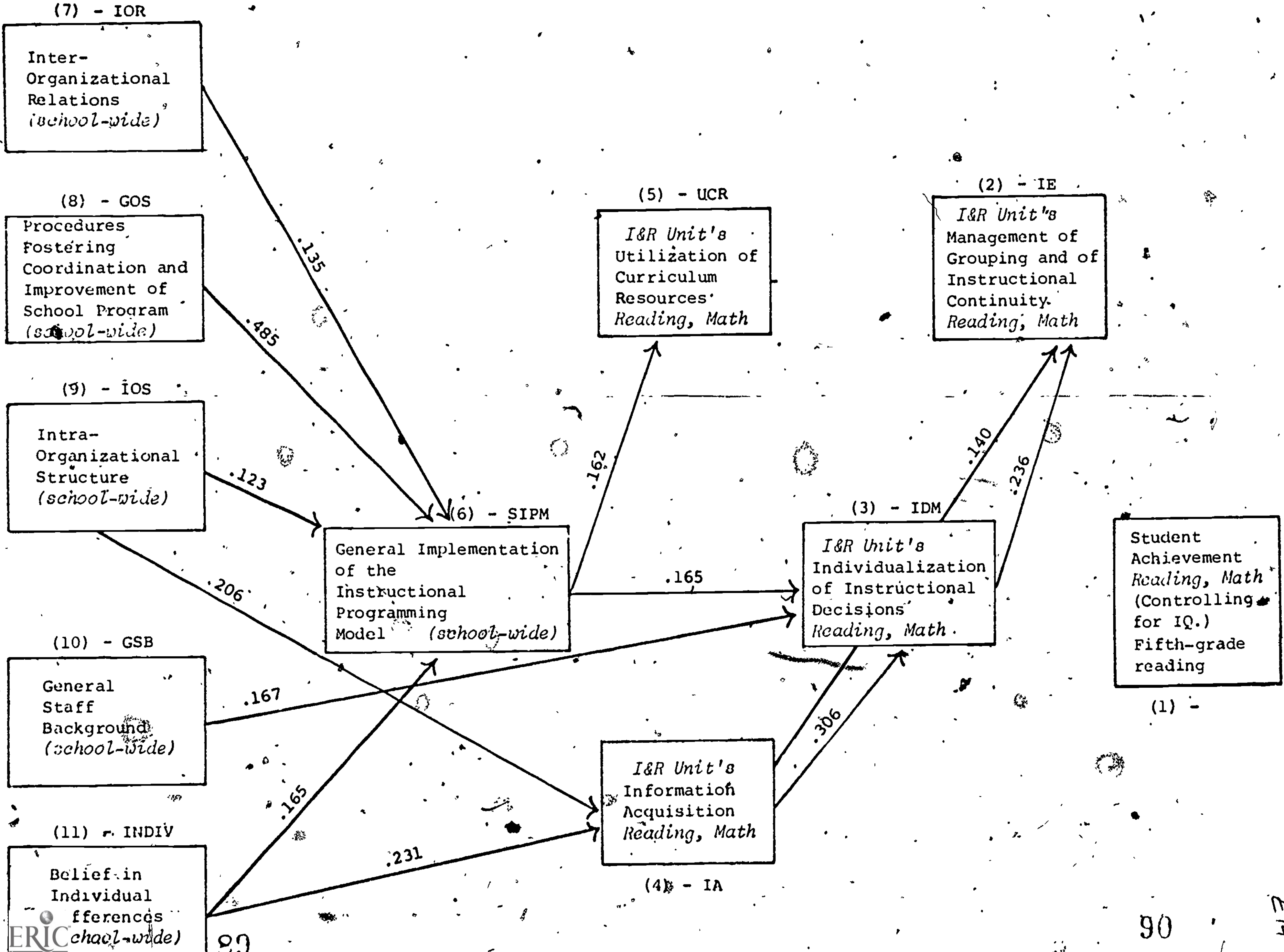


Figure 7. Fitted model for second-grade reading.



E14

Figure 8. Fitted model for fifth-grade reading.

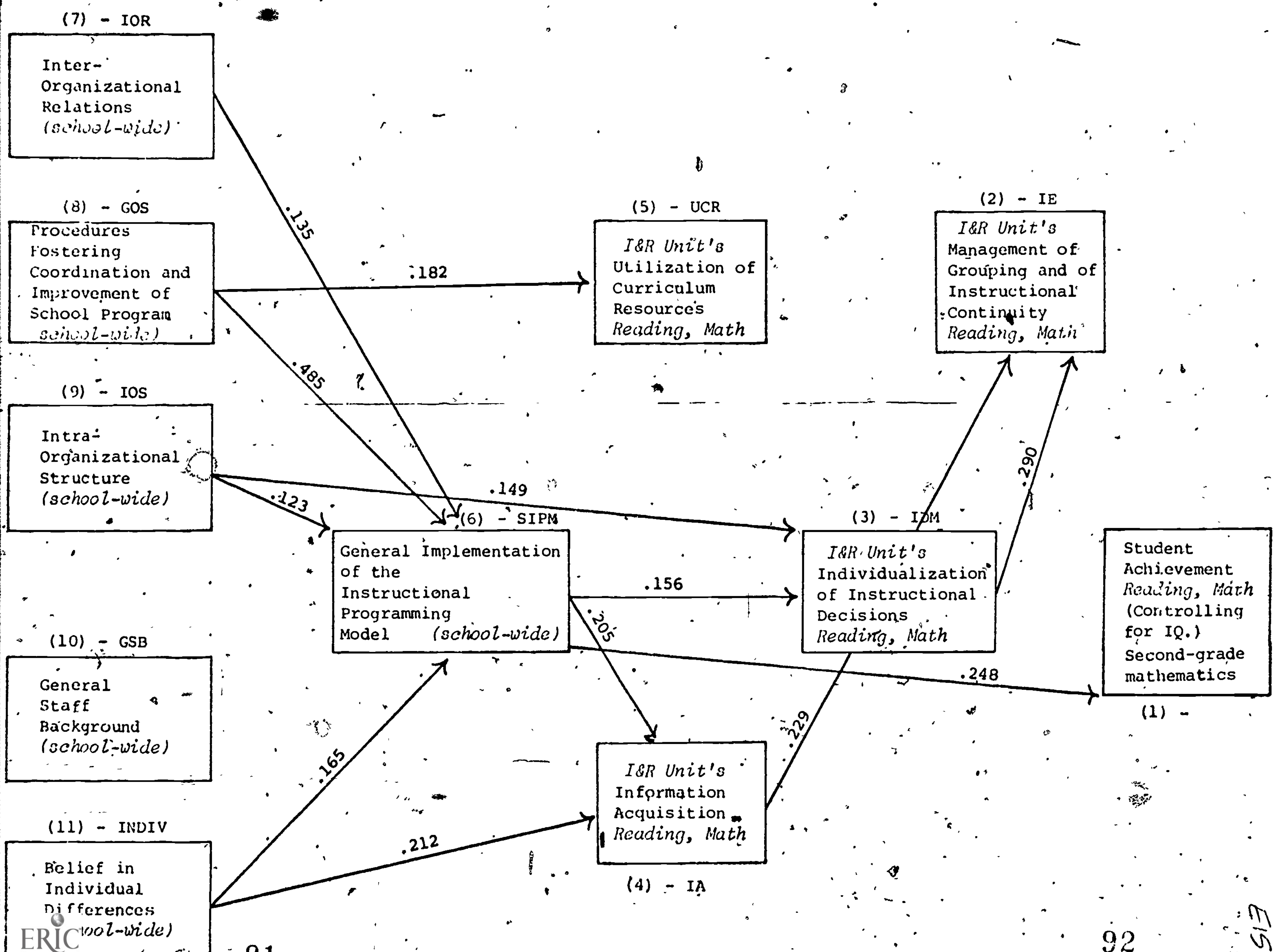


Figure 9. Fitted model for second-grade mathematics.

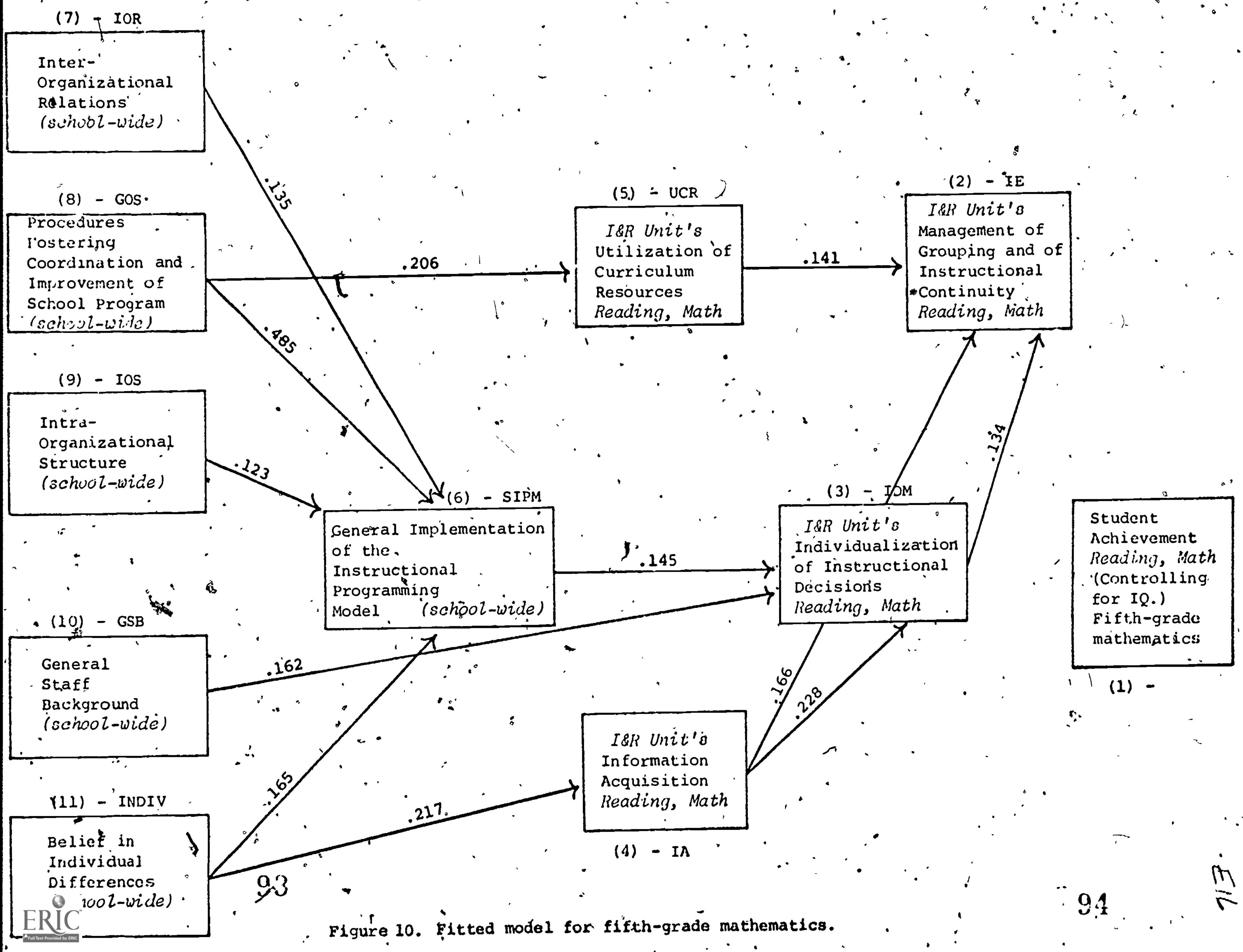


Figure 10. Fitted model for fifth-grade mathematics.

Orthogonalization of organizational variables. The six organizational variables in the model were transformed through principal components factor analysis into a set of six uncorrelated (orthogonal) factors. Varimax rotation was used on the full set of principal components. The net effect of performing these transformations was to impose orthogonality on the set of six predictors, while preserving as closely as possible a one-to-one correspondence between variables and principal components. Correlations between the six orthogonal factors and the six untransformed variables are given in Table 13; these correlations are a special type of factor loading. As is evident in Table 13, each of the six factors is a good proxy for one of the variables and is negligibly related to the other five variables. Thus, two orthogonalized predictors are substantively meaningful, so their coefficients in a regression equation do permit some interpretation.

When predictors are orthogonal like this, standardized regression coefficients are equivalent to correlation coefficients. Therefore, results based on orthogonalized predictors are presented as correlation coefficients.

Orthogonalization of instructional variables. For each of the four models (grade 2 reading, grade 5 reading, grade 2 math, grade 5 math), the four instructional variables in the model were transformed through principal components factor analysis into a set of four uncorrelated (orthogonal) factors. Correlations between the untransformed variables of each model and the corresponding orthogonal factors are given in Table 14.

Correlations between the organizational factors of Table 13 and the instructional factors of the respective models are given in Table 15. As can be seen from this table, the correlations between the organizational and instructional factors are generally nonsignificant; even those that are

Table 13

Correlations Between the Organizational Variables and their Varimax-rotated Principal Components

	Factors					
Variables	GOSFAC	IORFAC	IPMFAC	INDIVFAC	IOSFAC	GSBFAC
GOS	.76	.38	.36	.23	.29	.10
IOR	.24	.89	.25	.20	.20	.09
IPM	.25	.26	.87	.22	.23	.12
INDIV	.14	.17	.17	.96	.06	.06
IOS	.06	.18	.19	.06	.94	.13
GSB	.06	.08	.09	.06	.11	.98
Eigenvalue of factor	.72	1.07	1.02	1.07	1.08	1.01

Table 14

Correlations Between Instructional Feature Variables

and their Varimax-rotated Principal Components

Variables	Factors			
	IEFAC	UCRFAC	IAFAC	IDMFAC
Grade 2 Reading				
IE	.98	.05	.06	.17
UCR	.04	.99	.07	.10
IA	.06	.07	.99	.09
IDM	.18	.11	.10	.97
Eigenvalue of factor	1.00	1.00	1.00	.99
Grade 5 Reading				
UCR	.99	.07	.05	.08
IE	.07	.98	.09	.13
IA	.05	.10	.98	.17
IDM	.08	.13	.17	.97
Eigenvalue of factor	.99	.99	1.00	.99

Variables	Factors			
	IEFAC	UCRFAC	IAFAC	IDMFAC
Grade 2 Mathematics				
IE	.98	.00	.14	.16
UCR	.00	1.00	.06	.03
IA	.13	.06	.99	.07
IDM	.16	.03	.07	.98
Eigenvalue of factor	1.00	1.00	1.00	1.00
Grade 5 Mathematics				
UCR	.99	.09	.06	.06
IE	.08	.99	.10	.09
IA	.06	.10	.99	.12
IDM	.06	.09	.12	.99
Eigenvalue of factor	1.00	1.00	1.00	1.00

Table 15

Correlations of Organizational and Instructional Factors

Organizational	Instructional UCRFAC	IEFAC	IAFAC	IDMFAC
Grade 2 Reading				
IOSFAC	.04	-.02	.01	.14*
IORFAC	.07	.00	.00	.05
INDIVFAC	.07	.13	.23*	-.08
SIPMFAC	.12	.14*	.22**	.15*
GSBFAC	.13	.00	.01	-.02
GOSFAC	.04	.00	.11	.16*
Grade 5 Reading				
IOSFAC	-.04	-.08	.20**	.22**
IORFAC	.10	.12	.04	.05
INDIVFAC	.09	.11	.20**	.17*
SIPMFAC	.13*	.08	.06	.15*
GSBFAC	.10	-.03	-.02	.18*
GOSFAC	.08	-.16	-.01	-.11

Organizational	Instructional UCRFAC	IEFAC	IAFAC	IDMFAC
Grade 2 Mathematics				
IOSFAC	-.01	.06	.06	.19**
IORFAC	-.04	-.05	.03	-.05
INDIVFAC	.11	.06	.24**	.06
SIPMFAC	.05	.11	.19**	.13*
GBSFAC	.06	.01	.03	.04
GOSFAC	.17*	.00	.10	.15*
Grade 5 Mathematics				
IOSFAC	-.02	-.03	.10	.15*
IORFAC	.03	.09	.10	-.01
INDIVFAC	.11	-.03	.18*	.13
SIPMFAC	.09	.02	-.04	.17*
GSBFAC	.01	-.08	-.02	.17*
GOSFAC	.17*	-.06	.07	-.07

*$p < .05$

**$p < .01$

statistically significant are nevertheless weak. Our discussion is limited to those correlations that were statistically significant in both grades 2 and 5 in either reading or math. Those relationships are listed in Table 16. Weak but statistically significant correlations occurring in only one grade were regarded as undeserving of serious attention.

The correlations in Table 16 are weak, but they are statistically significant. Except for the relationship between GOSFAC and UCRFAC, which was statistically significant at both grades in mathematics only, the other three relationships are statistically significant in both reading and mathematics at both grades 2 and 5. Thus, although those correlations are weak, they do persist at a statistically significant level in four out of four circumstances.

Two of the persistently significant correlations involve the orthogonalized measure (IDMFAC) of the I & R unit's Individualization of Instructional Decisions (IDM). I & R units that scored high on IDM tended to be part of a school that scored high on schoolwide aspects of the implementation of the Instructional Programming Model (IPM), as indicated by the persistent correlation between SIPMFAC and IDMFAC. This finding is consistent with the assumption made in IGE that schoolwide implementation of the IPM affects its implementation at the level of individual I & R units. I & R units that scored high on IDM also tended to be part of a school that conformed to the intra-organizational structure recommended for IGE schools, as indicated by the persistent correlation between IOSFAC and IDMFAC.

The other relationship marked by persistently significant correlations is that between INDIV (a measure of the extent to which teachers in a school believe that individual differences are important to consider when making

Table 16

Bivariate Relationships that had Statistically Significant Correlations in Both Grades

	Correlations	
	Grade 2	Grade 5
Reading		
IOSFAC & IDMFAC	.14	.22
INDIVFAC & IAFAC	.23	.20
SIPMFAC & IDMFAC	.15	.15
Mathematics		
IOSFAC & IDMFAC	.19	.15
INDIVFAC & IAFAC	.24	.18
SIPMFAC & IDMFAC	.13	.17
GOSFAC & UCRFAC	.17	.17

instructional decisions) and IA (a measure of the extent to which an I & R unit collects information about individual differences in reading and mathematics, respectively). The correlations supporting this statement are those between the orthogonalized measures INDIVFAC and IAFAC. Although these correlations are weak, they are persistently significant, and they do suggest that teachers who believe in the instructional importance of individual differences are more likely to collect information of a kind that will support the individualization of instructional decisions.

As mentioned earlier, the correlation between GOSFAC and UCRFAC was significant at both grades 2 and 5 only in mathematics. The a priori model did predict positive correlation between GOS (Procedures Fostering Coordination and Improvement of the School Program) and UCR (Utilization of IPM-Compatible Curriculum Resources by I & R Units), but the model predicted the positive correlation in both reading and mathematics, not just in mathematics.

Job Satisfaction

The premise that teacher job satisfaction is affected by particular organizational features of IGE has been represented as a network of postulated causal links among the variables. Figure 11 presents the various causal links in diagram form. It shows the paths of influence assumed a priori to underlie the relationships between teacher job satisfaction and the other variables.

Four variables, Interorganizational Relations (IOR), Procedures Fostering Coordination and Improvement of the School Program (GOS), Intraorganizational Structure (IOS), and General Implementation of the Instructional Programming Model (SIPM), are hypothesized to relate to teacher job satisfaction. The model in Figure 11 suggests that the amount of communication with other IGE

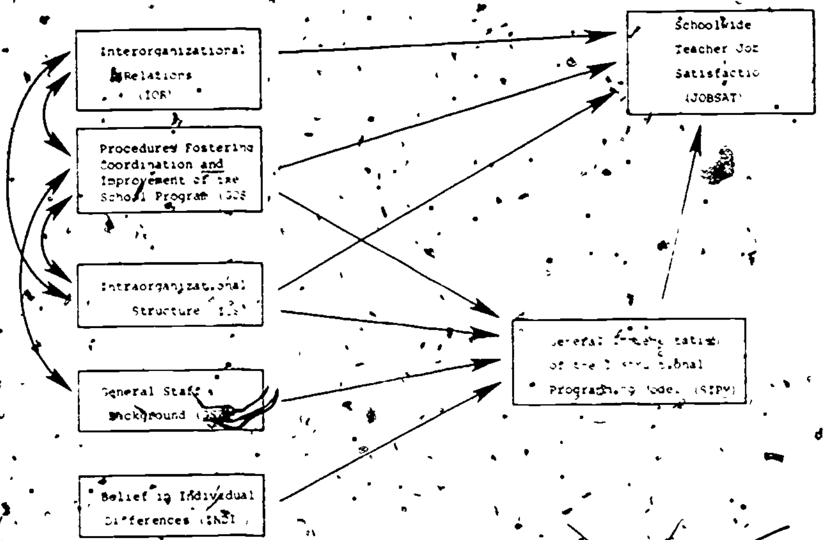


Figure 11. Phase I structural model predicting teacher job satisfaction.

schools (IOR) should help teachers feel better about their skills and be more satisfied with teaching. The structural model also suggests that schools with a well-functioning IGE program (schools high on GOS, IOS, and SIPM) would have satisfied teachers. The model indicates that General Staff Background (GSB) and Belief in Individual Differences (INDIV) indirectly affect teacher job satisfaction by way of SIPM. Factors affecting SIPM are discussed more completely in Technical Reports 510 and 511 (Price, Janicki, Van Deventer, & Romberg, 1980a, 1980b).

As with the analyses presented already, the analysis began with an estimation of the a priori model. The correlations of schoolwide variables with the job satisfaction variable are given in Table 17. Each endogenous variable was regressed on its theorized causes. Figure 12 shows the coefficients as estimated for the a priori model for teacher job satisfaction.

As a set of predictors, organizational features in the a priori model account for 29.1% of the variance in teacher job satisfaction, which is statistically significant ($p < .001$). For studies of this kind, 29.1% of the variance is good prediction. Whether one should regard good empirical prediction like this as practically significant depends on the interpretation given to the empirical relations, however strong they are. By adding to the prediction equation exogenous variables that had not been included a priori as predictors of teacher job satisfaction, the multiple R^2 is .301, which is only a 1% increase in variance explained. That small increase is not statistically significant ($p < .25$). Since the a priori model is approximately as predictive as is the fuller model, we are inclined to judge the a priori model to be adequate.

Table 17

Correlations of Schoolwide Variables with

the Job Satisfaction Variable

Schoolwide variables	JOBSAT
SIPM	.41421
IOR	.45554
GOS	.49912
IOS	.18243
GSB	.04655
INDIV	.32409

However, the negative coefficient of IOS does give some pause; the a priori model would have that coefficient be positive. The zero-order correlation of IOS with JOBSAT is, in fact, positive. It is .18 ($p < .05$, one-tailed test). The negative coefficient is inconsistent with the interpretation implicit in the a priori model. The contradictory interpretation is that the features measured by IOS have a weakly negative effect on JOBSAT, which is not apparent by examination of zero-order correlations because of the countervailing influence of IOR, GOS, and SIPM, with which IOS is positively correlated. However, neither the a priori interpretation nor the contradictory interpretation are unequivocally supported by these results.

As noted in earlier analyses, the intercorrelation of predictors makes any interpretation of particular coefficients treacherous. For that reason, the procedure of orthogonalizing predictors was again used. Variables causally prior to JOBSAT in the model were transformed through principal components factor analysis into a set of uncorrelated (orthogonal) factors. This procedure, its requirements, and its limitations were described earlier.

The uncorrelated, substantively interpretable factors defined in Table 13 were then used as predictors of JOBSAT. By having predictors that are uncorrelated among themselves, path coefficients can be compared and tested without ambiguity--as long as the variables used as predictors are themselves substantively meaningful.

Organizational features as predictors. Factors corresponding to three organizational features had correlations with JOBSAT that were statistically significant and of noteworthy magnitude. The factor corresponding to GOS had

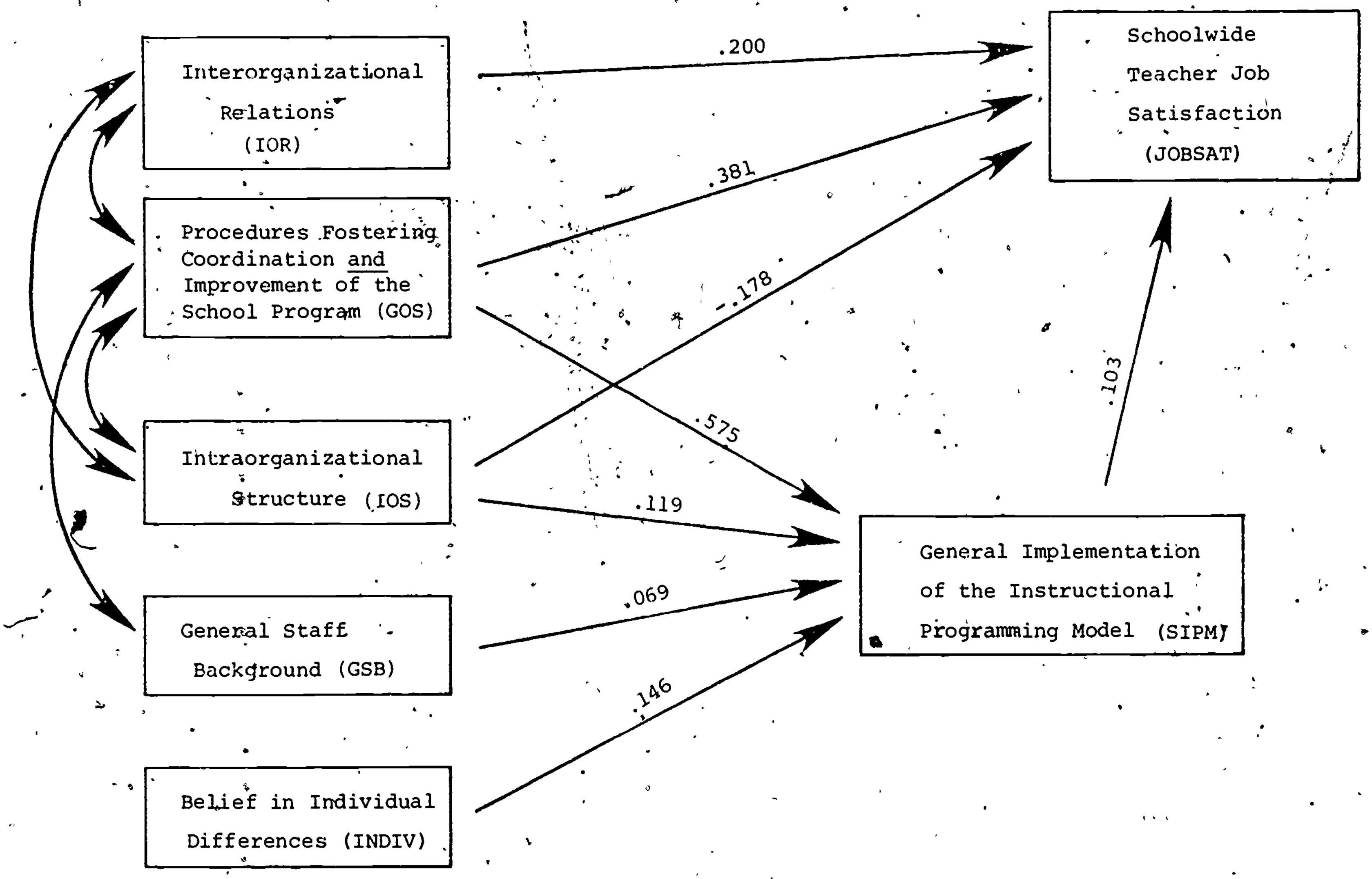


Figure 12. Estimated model for teacher job satisfaction.

a correlation of .32 ($p < .0001$, one-tailed test) with JOBSAT. The factor corresponding to IOR had a correlation of .31 ($p < .0001$, one-tailed test), and the factor corresponding to SIPM had .24 ($p < .01$, one-tailed test). The zero-order correlation coefficients reported here are equivalent to the standardized regression coefficients of JOBSAT regressed on the factors corresponding to GOS, IOR, and SIPM ($R^2 = .261$). That equivalence is inherent to perfectly uncorrelated predictors. The zero-order correlation of the factor corresponding to IOS, another organizational variable, is .02, which is statistically nonsignificant and trivially small.

School staff measures as predictors. Besides the four organizational features discussed in the preceding paragraph, two measures of the school staff, INDIV and GSB, were treated as causally prior to JOBSAT. The correlation of the factor corresponding to INDIV with JOBSAT is .20, which is statistically significant ($p < .01$, one-tailed test). The correlation of the factor corresponding to GSB with JOBSAT, on the other hand, is -.03, which is trivially small and statistically nonsignificant. The addition of the factor corresponding to INDIV to the predictor set of GOS, IOR, and SIPM improved the prediction of JOBSAT from $R^2 = .261$ to $R^2 = .299$, a statistically significant increase.

Conclusion

The conclusion of this report considers the findings of the Phase I study, dividing them for sake of exposition into five parts:

1. Relationships between organizational features and instructional practices of IGE schools.

2. Relationships between staff beliefs and instructional practices.

3. Relationships between instructional practices of IGE schools and measures of students' achievement.

4. Relationships between organizational features of IGE schools and students' achievement.

5. Correlates of teachers' job satisfaction in IGE schools.

Organizational Features and Instructional Practices

Part of IGE Schooling consists of organizational features designed to facilitate the instructional practices that compose the Instructional Programming Model. Because those organizational features have as a primary purpose the facilitation of certain instructional practices, the Phase I study examined empirically the implied relations between organizational features and instructional practices. Some implied correlations between organizational features and instructional practices were borne out in the Phase I data; others were not. The implied relations that were borne out empirically were presented in Table 15. The practice of individualizing instructional decisions--an instructional practice pursued to varying degrees by I & R units in the schools studied--does seem to be facilitated

by certain schoolwide organizational features, in particular the schoolwide implementation of the Instructional Programming Model and the intraorganizational structure of the schools.

Schoolwide implementation of the IPM. The extent to which the Instructional Programming Model (IPM) had been implemented by the school in general (and not simply in the I & R units studied) was positively correlated with the degree to which the specific I & R units under study engaged in the individualization of instructional decisions. This finding may surprise no one, but it does attest to the susceptibility of small groups of teachers to the larger (schoolwide) milieu in which they are situated.

Intraorganizational structure. Another aspect of the larger milieu that was consistently associated with I & R units' individualization of instructional decisions is the intraorganizational structure of the school. (This is a proxy for various structural arrangements distinctive to IGE schools, such as the organization of staff and students into I & R units and the existence of a functioning Instructional Improvement Committee.) These correlations with I & R units' individualization of instructional decisions offer some vindication of the organizational theory contained in the IGE system. According to that theory, individuals who are part of an organization are affected by controllable features of that organization.

Implied correlations not found. Some implied correlations between organizational features and instructional practices were not borne out by the data in a consistent fashion. They were:

1. an expected connection between the interorganizational relations of a school and the utilization of IPM-compatible curriculum materials by I & R units in that school.

2. expected connections between schoolwide implementation of the IPM and

   a. utilization of IPM-compatible materials by I & R units,

   b. collection of information about individual differences,

   c. the I & R unit's management of grouping and instructional continuity.

In cases such as these, where expected relations were not found, three types of explanations can be offered. The first might be called a "model-blaming" explanation, because it faults the underlying model that has failed under test. The second and third might be called "test-blaming," because they fault the procedures that have been used to perform the test of the model. The second attributes the lack of observed relation to faulty measurement of the predictor variables (in this case, measures of schoolwide organizational features). This explanation seems to have little ground in this case for a couple of reasons. As noted earlier in this report, the Phase I measures of schoolwide organizational features agreed reasonably well with counterpart measures obtained in the Phase II field validation study. Furthermore, other parts of the model were borne out by correlations involving these variables, a circumstance which should not have arisen if these variables were badly measured.

The third type of explanation attributes the lack of observed relation to faulty measurement of the predicted variables (in this case, measures of I & R units' instructional practices). This type of "test-blaming" explanation cannot be seriously disputed. There was no field validation of these measures, there was abundant opportunity for distorted information to enter

the questionnaires, and there are no redeemingly high correlations to suggest that these variables were measured reliably. The measures of I & R units' instructional practices are the least trustworthy part of the Phase I study. For this reason, we are disinclined to engage in "model-blaming" when data involving these particular measures are involved.

Staff Beliefs and Instructional Practices

The beliefs that staff members hold about the value of IGE are obviously important. The instructional practice of collecting information about individual differences between students in content areas (reading and mathematics) was, as expected, correlated with a measure of the extent to which teachers in a school believe that individual differences are important to consider when making instructional decisions. In the presence of other factors likely to affect this important element of IGE practice, it is noteworthy that its strongest predictor in Phase I was the extent to which teachers in a school believed in what one could reasonably argue is the most basic tenet of the Instructional Programming Model; namely, that individual differences are pertinent to instructional decisions.

Instructional Practices and Student Achievement

In no instance--not in Reading, not in Math, not in Grade 2, not in Grade 5--was there a statistically significant correlation between a measure of instructional practices and a measure of student achievement. Besides the "model-blaming" explanation for this, which would fault the Instructional Programming Model, there are two other types of explanation. One attributes

the weak relations to bad measurement of student achievement; the other, to bad measurement of instructional practices. Student achievement, we believe, was measured reliably. Despite the reliability with which student achievement was measured, any standardized, group-administered test can be criticized as an imperfect reflection of what children know about the area assessed by the test. Persons who wish to make that criticism of the student achievement measures--a criticism of construct validity--must concomitantly dismiss any favorable findings based on outcome measures such as these. We have already mentioned the low trust we place in our measures of instructional practices. Consequently, unreliability in those measures may have attenuated correlations between instructional practices and student achievement. For that reason, we are disinclined to use this particular negative finding as a basis for "model-blaming."

Organizational Features and Student Achievement

Expected correlations between organizational features and student achievement were not found, despite reliable measurement of both classes of variables. All measures of organizational features were trivially and nonsignificantly associated with student achievement measures. With regard to this negative finding, a "model-blaming" explanation is the most plausible. Specifically, these findings indicate that implementation of the surface organizational features with which IGE is commonly identified offer no assurance at all the student achievement will be raised. Evidently, the instructional effectiveness of I & R units included in Phase I did not depend on the degree to which IGE organizational features had been implemented.

in the school of which they were part. This finding (based on questionnaire data) resonates with the findings of other phases of the IGE Evaluation, all of which have converged on the conclusion that surface orthodoxy reveals little if anything about the value of an educational program. Nor does it reveal whether the Instructional Programming Model is practiced in a form that would be recognizable by its developers.

Correlates of Teachers' Job Satisfaction

The foregoing discussion followed a long-standing tradition in educational evaluation by gauging the worth of an educational option on the basis of its effects on the students' achievement. There do exist other grounds on which to evaluate educational programs. In an era when "teacher burnout" has become a household word, one evident alternative is to evaluate the effects of programs on staff morale. On those grounds, three organizational features commonly associated with IGE fare well. Three schoolwide organizational features have positive correlations (and positive path coefficients) with the schoolwide measure of teacher job satisfaction. Those three features are: (1) the interorganizational relations of the school, (2) the existence of procedures fostering coordination and improvement of the school program, and (3) general, schoolwide implementation of the Instructional Programming Model. Underlying these positive relations appear to be two factors. One is that teachers like the contact with other adults provided by IGE, especially the contact with a professional network that extends beyond their particular school. The other is that job satisfaction is derived from a belief that their instructional efforts are effective, a belief commonly held by teachers of schools in which the IPM had been implemented on a schoolwide basis.

References

Anscombe, F. J. Rejection of ouliers. Technometrics, 1960, 2, 123-147.

Barth, R. S. So you want to change to an open classroom. Phi Delta Kappan, 1971, 97-99.

Berman, P., & McLaughlin, M. E. Implementation of educational innovation. Educational Forum, 1976, 345-360.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. P., & York, R. L. Equality of educational opportunity. Washington, D.C.: U.S. Government Printing Office, 1966.

Cronback, L. J., & Meehl, P. E. Construct validity in psychological tests. Psychological Bulletin, 1955, 52, 281-302.

DeValut, M. V. Descriptor for individualized instruction. Madison: Wisconsin Center for the Analysis of Individualized Instruction, 1973.

Dixon, W. J., & Tukey, J. W. Approximate behavior of the distribution of winsorized t (trimming/winsorization 2). Technometrics, 1968, 10, 83-98.

Firkowska, A., Ostrowska, A., Sokolowska, M., Stein, Z., Susser, M., & Wald, I. Cognitive development and social policy. Science, 1978, 200, 1357-1362.

Goldberger, A. Econometric theory. New York: Wiley, 1964.

Gordon, R. A. Issues in multiple regression. American Journal of Sociology, 1968, 77, 592-616.

Harman, H. Modern factor analysis. Chicago: University of Chicago Press, 1967.

Harris, M. L., & Harris, C. W. A structure of concept attainment abilities. Madison: Wisconsin Research and Development Center for Individualized Schooling, 1973.

Ironside, R. A. The 1971-72 nationwide installation of the multiunit/IGE model for elementary schools: A process evaluation. (OEC-0-71-3705) Princeton: Educational Testing Service, 1972. ED 066 623.

Ironside, R. A., & Conaway, L. IGE evaluation, Phase II on-site validation and descriptive study. Final report (Technical Report No. 499). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1979.

Klausmeier, H. J. Individually guided education in elementary and middle schools: A handbook for implementors and college instructors (Preliminary edition). Madison: IGE Teacher Education Project, University of Wisconsin-Madison, 1976.

Klausmeier, H. J. Origin and overview of IGE. In H. J. Klausmeier, R. A. Rossmiller, & M. Saily (Eds.), Individually guided elementary education: Concepts and practices. New York: Academic Press, 1977.

Klausmeier, H. J., Karges, M. L., & Krupa, W. E. Outcomes of IGE and performance objectives for implementation. In H. J. Klausmeier, R. A. Rossmiller, & M. Saily (Eds.), Individually guided elementary education: Concepts and practices. New York: Academic Press, 1977.

Klausmeier, H. J., Morrow, R., & Walter, J. E. Individually guided education: Guidelines for implementation. Madison: Wisconsin Research and Development Center for Individualized Schooling, 1968.

Klausmeier, H. J., Rossmiller, R. A., & Saily, M. (Eds.). Individually guided elementary education: Concepts and practices. New York: Academic Press, 1977.

Lipham, J. M., & Fruth, M. J. The principal and individually guided education. Reading: Addison-Wesley, 1976.

Mendenhall, D. R. Relationship of organizational structure and leadership behavior to staff satisfaction in IGE schools (Technical Report No. 412). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1977.

Otto, W. The Wisconsin Design: A reading program for individually guided education. In H. J. Klausmeier, R. A. Rossmiller, & M. Saily (Eds.), Individually guided elementary education: Concepts and practices. New York: Academic Press, 1977.

Price, G. G. Sampling for Phase I of the IGE evaluation (Working Paper No. 223). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1977.

Price, G. G., Janicki, T. C., Howard, J. A., Stewart, D. M., Buchanan, A. E., & Romberg, T. A. Procedures used in Phase I of the IGE evaluation to scale the variable Interorganizational Relations (Technical Report No. 476). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1978. (a)

Price, G. G., Janicki, T. C., Howard, J. A., Stewart, D. M., Buchanan, A. E., & Romberg, T. A. Procedures used in Phase I of the IGE evaluation to scale the variable Procedures Fostering Coordination and Improvement of the School Program (Technical Report No. 477). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1978. (b)

Price, G. G., Janicki, T. C., Howard, J. A., Stewart, D. M., Buchanan, A. E., & Romberg, T. A. Procedures used in Phase I of the IGE evaluation to scale the variable General Staff Background (Technical Report No. 478). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1979. (c)

Price, G. G., Janicki, T. C., Howard, J. A., Stewart, D. M., Buchanan, A. E., & Romberg, T. A. Procedures used in Phase I of the IGE evaluation to scale the variable Intraorganizational Structure (Technical Report No. 479). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1978. (d).

Price, G. G., Janicki, T. C., Howard, J. A., Stewart, D. M., Buchanan, A. E., & Romberg, T. A. Procedures used in Phase I of the IGE evaluation to scale the variable Belief in Individual Differences (Technical Report No. 480). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1978. (e)

Price, G. G., Janicki, T. C., Howard, J. A., Stewart, D. M., Buchanan, A. E., & Romberg, T. A. Procedures used in Phase I of the IGE evaluation to scale the variable School's Demographic Setting (Technical Report No. 482). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1978. (f)

Price, G. G., Janicki, T. C., Howard, J. A., Stewart, D. M., Buchanan, A. E., & Romberg, T. A. Procedures used in Phase I of the IGE evaluation to scale the variable General Implementation of the Instructional Programming Model (Technical Report No. 483). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1978. (g)

Price, G. G., Janicki, T. C., Howard, J. A., Stewart, D. M., Buchanan, A. E., & Romberg, T. A. Procedures used in Phase I of the IGE evaluation to scale the variable Teacher Job Satisfaction (Technical Report No. 484). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1978. (h)

Price, G. G., Janicki, T. C., Howard, J. A., Stewart, D. M., Buchanan, A. E., & Romberg, T. A. Procedures used in Phase I of the IGE evaluation to scale the variable Utilization of Curriculum Resources (Technical Report No. 485). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1978. (i)

Price, G. G., Janicki, T. C., Howard, J. A., Stewart, D. M., Buchanan, A. E., & Romberg, T. A. Procedures used in Phase I of the IGE evaluation to scale the variable Information Acquisition (Technical Report No. 486). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1978. (j)

Price, G. G., Janicki, T. C., Howard, J. A., Stewart, D. M., Buchanan, A. E., & Romberg, T. A. Procedures used in Phase I of the IGE evaluation to scale the variable Individualization of Instructional Decisions (Technical Report No. 487). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1978. (k)

Price, G. G., Janicki, T. C., Howard, J. A., Stewart, D. M., Buchanan, A. E., & Romberg, T. A. Procedures used in Phase I of the IGE evaluation to scale the variable Management of Grouping and of Instructional Continuity (Technical Report No. 489). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1978. (l)

Price, G. G., Janicki, T. C., VanDeventer, H. C., Jr., & Romberg, T. A. Organizational features of IGE schools as correlates of teacher job satisfaction (Technical Report No. 512). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1979.

Price, G. G., Janicki, T. C., VanDeventer, H. C., Jr., & Romberg, T. A. Organizational and instructional features that affect grade 2 and grade 5

reading achievement in IGE schools: A causal model analysis (Technical Report No. 510). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1980. (a)

Price, G. G., Janicki, T. C., VanDeventer, H. C., Jr., & Romberg, T. A. Organizational and instructional features that affect grade 2 and grade 5 math achievement in IGE Schools: A causal model analysis (Technical Report No. 511). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1980. (b)

Price, G. G., VanDeventer, H. C., Jr., Janicki, T. C., & Romberg, T. A. Procedures used in Phase I of the IGE evaluation to scale, adjust, and aggregate reading and math achievement in grade 2 (Technical Report No. 508). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1979. (a)

Price, G. G., VanDeventer, H. C., Jr., Janicki, T. C., & Romberg, T. A. Procedures used in Phase I of the IGE evaluation to scale, adjust, and aggregate reading and math achievement in grade 5 (Technical Report No. 509). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1979. (b)

Romberg, T. A. IGE evaluation: Perspectives and a plan (Working Paper No. 183). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1976.

Romberg, T. A. Developing mathematical processes: The elementary mathematics program for individually guided education. In H. J. Klausmeier, R. A. Rossmiller, & M. Saily (Eds.), Individually Guided elementary education: Concepts and practices. New York: Academic Press, 1977.

Rossmiller, R. A., & Geske, T. G. Economic analysis of education: A conceptual framework (Theoretical Paper No. 68). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1977.

Stewart, D. M. Instrumentation for Phase I of the IGE evaluation. (Unpublished Project Paper 77-1). Madison: Wisconsin Research and Development Center for Individualized Schooling, 1977.

Venezky, R. L., & Pittelman, S. D. PRS: A pre-reading skills program for individually guided education. In H. J. Klausmeier, R. A. Rossmiller, & M. Saily (Eds.), Individually guided education: Concepts and practices. New York: Academic Press, 1977.

ASSOCIATED FACULTY

Thomas P. Carpenter
Professor
Curriculum and Instruction

W. Patrick Dickson
Assistant Professor
Child and Family Studies

Fred N. Finley
Assistant Professor
Curriculum and Instruction

Lloyd E. Frohreich
Professor
Educational Administration

Maureen T. Hallinan
Professor
Sociology

Dale D. Johnson
Professor
Curriculum and Instruction

Herbert J. Klausmeier
V. A. C. Henmon Professor
Educational Psychology

Joel R. Levin
Professor
Educational Psychology

James M. Lipham
Professor
Educational Administration

Cora B. Marrett
Professor
Sociology and Afro-American Studies

Fred M. Newmann
Professor
Curriculum and Instruction

Wayne Otto
Professor
Curriculum and Instruction

Penelope L. Peterson
Associate Professor
Educational Psychology

W. Charles Read
Professor
English and Linguistics

Thomas A. Romberg
Professor
Curriculum and Instruction

Richard A. Rossmiller
Professor
Educational Administration

Peter A. Schreiber
Associate Professor
English and Linguistics

Ronald C. Serlin
Assistant Professor
Educational Psychology

Barbara J. Shade
Assistant Professor
Afro-American Studies

Marshall S. Smith
Center Director and Professor
Educational Policy Studies and Educational Psychology

Aage B. Sørensen
Professor
Sociology

James H. Stewart
Assistant Professor
Curriculum and Instruction

B. Robert Tabachnick
Professor
Curriculum and Instruction and Educational Policy Studies

Gary G. Wehlage
Professor
Curriculum and Instruction

Alex Cherry Wilkinson
Assistant Professor
Psychology

Louise Cherry Wilkinson
Associate Professor
Educational Psychology

Steven R. Yussen
Professor
Educational Psychology